UNCLASSIFIED

AD 439387

DEFENSE DOCUMENTATION CENTER

FOR

SCIENTIFIC AND TECHNICAL INFORMATION

CAMERON STATION, ALEXANDRIA, VIRGINIA

19990322091

UNCLASSIFIED

(1)

(6) FACTOR ANALYSIS OF DATA MATRICES,

PART IV,

(10) by Paul Horst,

~~Project Staff~~

~~Fern Thomas~~
~~Helen Ranck~~

(15)

~~This Study Was Supported~~ in Part by

~~Office of Naval Research~~ Contract Nonr-477(33) and

PHR ~~and~~

~~Public Health Research~~ Grant MH00743-08.

Principal Investigator: Paul Horst

(8)

(11) Feb~~ruary~~ 1964

(5) 370200

~~University of~~ Washington Univ.,
Seattle, ~~Washington~~

FACTOR ANALYSIS OF DATA MATRICES

FOREWORD

This is Part IV of a series of reports on rationales and techniques of matrix factoring which play an important role in multivariate analysis techniques. Indeed, it may well be said that all adequate models and methods of multivariate analysis are special cases of matrix factoring techniques. The more traditional methods of factor analysis, in particular, are special cases of more general matrix factoring techniques, as are also all multiple regression models.

TABLE OF CONTENTS

PART IV

CHAPTER 13

THE PROBLEM OF ORIGIN

13.1 Kinds of Origin Problems

In the factor analysis techniques previously discussed, we have expressly or implicitly assumed that we begin with either a correlation matrix or a matrix of standardized measures. In the latter case the means of the columns are 0 and their variances unity.

We have indicated in Chapter 4 that the results of a factor analysis are dependent on the selection of both scale and origin for each of the variables. We have pointed out that for some variables, such as height, weight, time, volume, and many physiological and other variables, there is a natural origin which means that when the value of the variable is 0, none of that attribute exists for a particular entity.

We have also indicated that in psychology and many other disciplines we may have measures in which it is difficult, if not impossible, to specify the true zero point for the particular attribute. For example, in psychological measures the score is frequently the number of items which are answered correctly. If the items are very difficult and are administered to a group of first grade children, perhaps none of the children will get any of the items correct. If these items are supposed to measure some kind of intelligence, we certainly may not assume that some of the children have no intelligence because they got none of the items correct.

The problem of determining the zero points for a particular group of measures on a number of entities is usually solved by subtracting the mean from each variable, so that the resulting measures indicate the deviation of each of the individuals from the mean of the group.

We also pointed out in Chapter 4 that the scale of one attribute may be quite different from that of another. For example, a difference between a score of 5 and a score of 10 on a test may not be comparable to the difference between scores of 5 and 10 on another test, because the one test may have many more items in it than the other, and the items in the former may be much more or much less difficult.

In this chapter, however, we shall not consider the problem of scaling the variables. It will be recalled that usually they are scaled so that the standard deviations are all equal to unity. The more general problem of scaling and its effect on factor analysis will be considered in Chapter 15. In this chapter we shall be concerned with problems of origin as they relate to factor analysis results.

We have in general three kinds of origin problems for a data matrix of raw measures. The first of these we may call the major transformation problem, the second, the minor transformation problem, and the third, the double transformation problem.

13.1.1 The Major Transformation Problem. If we have a vertical data matrix having more entities than attributes, the major transformation procedure consists of premultiplying the raw data matrix on the left by a matrix whose order is equal to the major dimension of the vertical matrix. The typical major transformation procedure for a data matrix consists simply of subtracting the mean of each column vector from every element in the column.

We can see, however, that this amounts to premultiplying the matrix by a special kind of matrix, as indicated in Eq. (13.1.1).

$$x = \left(I - \frac{1\,1'}{N}\right) X \qquad (13.1.1)$$

Here we use X on the right of the equation to represent the raw score matrix. The x on the left represents the deviation score matrix. The order of the unit vector is N. The matrix in parentheses may be called a centering matrix because its premultiplication into the raw score matrix results in the x matrix whose elements are centered by columns.

It is easy to see from Eq. (13.1.1) that if we premultiply both sides of this equation by the row unit vector, we must get a null row vector. This is because premultiplication of the matrix in parentheses by the unit vector yields the null vector, and therefore the left hand side of Eq. (13.1.1) must vanish.

We can show very simply that this matrix operation on the raw score matrix is the same as if we had subtracted the mean of each column of the raw score matrix from each element of the corresponding column. We indicate the computation of the vector of column means by Eq. (13.1.2).

$$M' = \frac{1' X}{N} \qquad (13.1.2)$$

This is the raw score matrix premultiplied by a row unit vector and divided by the number of rows. This is, of course, the conventional definition of a vector of means.

The subtraction of the mean of each column from each of the elements in that column is indicated by Eq. (13.1.3).

$$x = X - 1\,M' \qquad (13.1.3)$$

It can readily be seen, by multiplying out the right hand side of Eq. (13.1.1) and using Eq. (13.1.2), that Eq. (13.1.3) results.

It may not always be true, however, that we wish to perform our analysis upon the deviation score matrix indicated in Eq. (13.1.3). Instead of subtracting the mean of each column from the raw score matrix, we may wish to add other constants to the columns.

This case is indicated by Eq. (13.1.4).

$$U = X - 1 V' \tag{13.1.4}$$

Here we may have a general vector of values, $\underline{V}$, in which each element in the $\underline{V}$ vector may be different. Some of the elements may be positive and some negative. If we have a set of measures taken from arbitrary origins, and if we have a good rationale for determining what the natural origins should be, we may adjust the matrix as in Eq. (13.1.4) so that the attributes of the matrix $\underline{U}$ on the left may be regarded as being measured from the natural origins of the attributes.

An example of natural origins occurs in learning data. One may, for example, have a group of subjects who are learning a skill such as typewriting. Measures of proficiency for a group of subject may be taken at weekly intervals. A speed score of 40 words per minute during the third week would be meaningful when compared with a speed score of 60 the fifth week. One may therefore construct a data matrix in which the rows are entities or persons, and the columns are scores made at successive time intervals. To take deviation measures for each of the successive time intervals rather than retain the original measures may lose precisely the information that one is interested in studying.

13.1.2 The Minor Transformation. We may have a type of origin problem less common than the one considered in the previous paragraphs. This may be

called a minor transformation. Here the data matrix is multiplied on the right by a centering matrix whose order is the minor order or width of the data matrix. Such a transformation is indicated in Eq. (13.1.5).

$$W = X\left(I - \frac{1\,1'}{n}\right) \qquad (13.1.5)$$

Here we have the raw score matrix $\underline{X}$, postmultiplied by the centering matrix in parentheses. This operation produces the matrix $\underline{W}$ on the left side of Eq. (13.1.5) whose rows sum up to 0. It is easy to see that if we postmultiply both sides of Eq. (13.1.5) by a column unit vector, the result must be a null vector. This is because the centering matrix postmultiplied by the unit vector yields a null vector.

Let us now see what this type of operation means. We indicate a column vector of row means by Eq. (13.1.6).

Let

$$m = \frac{X\,1}{n} \qquad (13.1.6)$$

We see that now the mean of each row is subtracted from each element in the corresponding row of the original matrix, as in Eq. (13.1.7).

$$W = X - m\,1' \qquad (13.1.7)$$

If we multiply out the right hand side of Eq. (13.1.5) and substitute from the left of Eq. (13.1.6), we get Eq. (13.1.7).

We may well ask why one should wish to center a matrix on the right in this fashion. One may have reason to believe that only the deviations of scores from a person's own mean are of significance. Certain models used in the measurement of personality traits result in right centered data matrices.

It may be that instead of subtracting the mean from each row of a data matrix, we wish to subtract some other value. For example, we may have a series of blood pressure readings over a period of time for a number of individuals. It may be thought that the significant measure is not the absolute blood pressure at any interval, but rather its deviation from the blood pressure during a given condition of the individual, such as when he is resting or when he first gets up in the morning.

In this case then, a different value may be subtracted from each of the row observations for the individuals, as indicated in Eq. (13.1.8).

$$W = X - v\,1' \qquad (13.1.8)$$

Here $\underline{v}$ is a vector which may represent some base state for each of the individuals, and its elements may vary from one individual to the next.

In any case, the results of the factor analysis will be influenced by the right centering type of operation or the generalization of it indicated in Eq. (13.1.8).

13.1.3 The Double Transformation. In the previous discussion of the minor transformation we assumed that no major or left transformation had been performed on the data matrix. Perhaps the more common case involving the minor transformation or right centering of the data matrix occurs when there has previously been a major or left centering transformation also. It is perhaps most common to center a data matrix on the right after it has been converted to a deviation score matrix whose column means are 0.

The case of the doubly centered matrix is indicated by Eq. (13.1.9).

$$z = \left(I - \frac{1\,1'}{N}\right) X \left(I - \frac{1\,1'}{n}\right) \qquad (13.1.9)$$

Here we have the raw score matrix premultiplied by a centering matrix and postmultiplied by another centering matrix. It must be observed, however, that these are not in general the same centering matrices, since the number of entities is not usually the same as the number of attributes. The number of entities is ordinarily greater than the number of attributes. Then the order of the left centering matrix will be larger than that of the right centering matrix. This is indicated by the scalar quantities used to divide the major product moment of the unit vectors in the parentheses. It will be noticed that on the left this is $\underline{N}$, indicating the number of cases, and on the right it is $\underline{n}$, indicating the number of variables.

Let us now define a matrix or scalar quantity, $\underline{\alpha}$, as in Eq. (13.1.10).

$$\alpha = \frac{1}{2}\,\frac{1'\,X\,1}{Nn} \tag{13.1.10}$$

This, we see, is obtained by taking the sum of all of the elements in the raw score matrix, and dividing it by the product of the number of entities by the number of attributes, and then taking one half this ratio. Obviously then, the value of $\underline{\alpha}$ is simply one half the average value of all of the elements in the data matrix.

We next define a row vector as in Eq. (13.1.11).

$$_M V' = M' \alpha 1' \tag{13.1.11}$$

We use a prescript $\underline{M}$ for the $\underline{V}$ vector, which is simply the $\underline{M}$ vector of Eq. (13.1.2) from each element of which has been subtracted the scalar $\underline{\alpha}$ given in Eq. (13.1.10).

We define another $\underline{V}$ vector with a prescript $\underline{m}$ as in Eq. (13.1.12).

$$_m V = m - \alpha 1 \tag{13.1.12}$$

This is obtained by subtracting from each element of the vector of row means defined in Eq. (13.1.6), the $\underline{\alpha}$ scalar calculated in Eq. (13.1.10).

We now define a $\underline{z}$ matrix as in Eq. (13.1.13).

$$z = X - 1\,{}_{M}V' - {}_{m}V\,1' \qquad (13.1.13)$$

This is obtained by subtracting from the raw score matrix two major products of vectors. The first of these is the unit vector postmultiplied by the vector calculated in Eq. (13.1.11), and the second is the vector calculated in Eq. (13.1.12) postmultiplied by the unit row vector. It can be proved from Eqs. (13.1.9) through (13.1.12) that this yields a matrix which is both left and right centered. This means that both the sums of rows and columns are 0.

It should be observed that the same results could have been obtained by first performing a left centering, as in Eq. (13.1.3), and then performing a right centering on the resulting matrix, as in Eq. (13.1.7). In this case, however, the right centering would have been performed on the $\underline{x}$ matrix calculated in Eq. (13.1.3), rather than on the raw score matrix calculated in Eq. (13.1.7).

As in the previous examples, one may also have a more rational basis for adding a particular scalar to each element in a column, and a particular scalar to each element in a row. This model is indicated in Eq. (13.1.14).

$$y = X - 1\,{}_{G}V' - {}_{G}v\,1' \qquad (13.1.14)$$

Here the $\underline{V}$ vector and the $\underline{v}$ vector in the terms on the right side of the equation consist of such elements.

In all of the three cases indicated, if the constants added produce right or left centering or both, the analysis is somewhat simpler than if general scalars are added, as in Eqs. (13.1.4), (13.1.8), and (13.1.14).

13.2 Ipsative Measures

We shall next consider a special type of data matrix which is a special case of the minor and double transformation types considered above. These are sometimes called ipsative measures.

There is a rather large class of problems in psychology which involves ipsative measures. These measures do not purport to be comparable from one individual to another for a given attribute, but only for different measures of the same individual. That is, the measures indicate the relative order of magnitude of the variables for a particular individual. For that individual the origin of the measures may be quite arbitrary. Measures such as these are known as ipsative, as distinguished from normative measures which do indicate for a particular variable the differences among individuals. This general problem has been extensively treated by Clemans (1956) in a doctor's dissertation.

First we shall consider how in psychology we may obtain matrices of data which are ipsative or which merely indicate from some arbitrary origin for each person his value on each of the variables in the set.

13.2.1 Definition of Ipsative Variables. We begin by giving a more general mathematical definition of ipsative variables than is commonly used. We have seen that by a right centering type of operation in which the row means are subtracted from each element in a row, we get a matrix the row sums of which are 0. We could now arbitrarily add some constant or scalar to each of the rows of this resulting matrix so that the rows would add up

to a constant instead of adding up to 0.

Next let us look at the problem in a different way. Suppose that we have a matrix of raw measures obtained in some particular manner, which we shall discuss in more detail later. Suppose, however, that this matrix exhibits the characteristic that the sum of rows all add up to the same constant, say, c, as indicated in Eq. (13.2.1).

$$X\,1 = c\,1 \tag{13.2.1}$$

This equation indicates that the data matrix postmultiplied by the unit vector is equal to some constant times the unit vector.

Suppose now we perform a left centering operation on this matrix, which amounts to putting it in deviation form by columns so that the sums of all columns are 0. This operation is indicated in Eq. (13.2.2).

$$x = \left(I - \frac{1\,1'}{N}\right) X \tag{13.2.2}$$

Recall that this operation simply subtracts the mean of each column from each element in that column.

We now prove that the resulting matrix x has the properties of a right centered matrix, as indicated in Eq. (13.2.3).

$$x\,1 = 0 \tag{13.2.3}$$

This means that the sums of rows of the resulting matrix are all 0.

The proof is as follows. From Eq. (13.2.2) we can write Eq. (13.2.4).

$$x\,1 = X\,1 - \frac{1\,(1'\,X\,1)}{N} \tag{13.2.4}$$

Here we simply expand the right side of Eq. (13.2.2).

Because of Eq. (13.2.1), we can write Eq. (13.2.5).

$$1' X 1 = c 1' 1 = c N \tag{13.2.5}$$

This means that the sum of the elements in the $\underline{X}$ matrix is the constant $\underline{c}$ times the number of entities, $\underline{N}$.

If we substitute Eqs. (13.2.1) and (13.2.5) in Eq. (13.2.4), we get Eq. (13.2.6).

$$x 1 = c 1 - c 1 = 0 \tag{13.2.6}$$

This proves that the sums of rows are 0 for the matrix in Eq. (13.2.2).

To state the case simply, if we have a matrix whose row elements add up to the same constant and we perform a left centering operation on this matrix, the resulting matrix is such that its row elements also add up to 0. While many of the matrices we deal with in psychology are not directly of the right centered type whose rows add up to 0, nevertheless, when they are left centered or put in deviation form, the rows are also in deviation form or have the property of right centered matrices.

13.2.2 Sources of Ipsative Variables. We shall now consider a number of different ways in which ipsative matrices can arise. We shall use the term ipsative to cover any set of measures with entity rows and column attributes such that the row sums add up to the constant, whether this constant is 0 or different from 0.

One type of model which involves the ipsative matrix is the differential prediction model. In this model we attempt to predict in which of a number of activity variables a person would be most likely to be superior. The question, however, is not what his score or performance would be in each

of the criterion variables as compared with other persons in the sample, but how good he would be in each criterion measure relative to his performance in each of the other criteria.

It can be shown in methods developed by Horst (1954) that a solution of such a problem results from the following operations. We require a matrix of intercorrelations of predictor with criterion variables. This may or may not be a matrix on which the correlations for all pairs of variables are based on all the cases. Usually it will not be, as we have seen in Chapter 11. In any case, suppose we have such a correlation matrix in which the rows are correlations of a given predictor variable with all the criterion variables, or in which a column is the correlation of a given criterion variable with all the predictor variables. If we take such a matrix and perform a right centering operation upon it, we then have a resulting matrix whose rows add up to 0. By means of methods which are beyond the scope of this text, we then use this right centered matrix together with other data to derive a matrix of prediction weights to apply to the predictor variables, so as to give the best differential prediction of success in the criterion variables.

Another source of ipsative measures comes from what is known as the forced choice type of psychological inventory. Here the subject is presented with pairs or groups of items. He is instructed to indicate which of these is most like him, which he most agrees with, or some other instruction which requires him to mark only one of each pair or group. If these scales are properly constructed, they have the advantage that the subject is choosing among the two or more items on the basis of two or more different traits, rather than on the basis of some single dimension, such as social

desirability. Examples of such inventories are the Edwards Personal Preference Schedule, the Kuder Preference Record, and others.

One of the characteristics of these measures is that normally the items are paired or grouped in such a fashion that the total of each person's score on each scale is a constant. That is, if the inventory is scores, say, for 15 different traits or measures, the sum of each person's measures will add up to the same constant. This means, then, that if a set of measures is transformed to deviation measures by a left centering operation, the sum of rows of the resulting matrix will be 0.

Another source of ipsatized measures may occur when it is desired to determine experimentally the properties of ipsative, as distinguished from normative, measures. Wright (1957) administered the Edwards Personal Preference Schedule to a group of individuals. The items in the schedule were also prepared in a rating scale format so that each item was presented singly, and the same subjects were requested to respond on a rating scale which indicated the extent to which the item applied to them. It was then possible to get a score for each of the traits presumably measured by the scale, by means of an appropriate scoring key. The problem was to see how the results of this kind of format compared with those of the forced choice, after the rating scale measures had been ipsatized by means of a right centering operation.

Another type of ipsatization seems to occur, in part, with self-appraisal inventories of the interest, personality, and temperament type, even though the items are not in paired or forced choice format. There is evidence to indicate that, if a person is presented with a set of items or statements in which he is asked to indicate how well each statement applies

to him, he engages in a sort of self-ipsatizing operation. In other words, he tends to adopt some sort of an average of all of the items as they apply to him, and then indicate which of the items is above and which below his own average for all of the items in the set. There seems to be a tendency to adjust oneself to the particular set in a relative rather than an absolute sense. This appears to be a special case of a general phenomenon in which it is easier to make comparative judgments rather than absolute judgments.

In any case, factor analysis of items of this kind seems to give results which indicate that a sort of partial ipsatization has been taking place, even though not a complete right centering type of operation.

13.2.3 Characteristics of Ipsative Matrices. The effect of either right or left centering of a matrix on the rank of this matrix is of considerable importance. It is especially so when ipsative measures are used as predictor variables. If, as is usually the case, there are more entities than attributes, then the ipsative matrix, after left centering or conversion to deviation form, is reduced to a rank one less than its width.

This can be seen from Eq. (13.1.5). We learned in Chapter 3 that the product of any two matrices cannot be greater than the rank of the factor of smaller rank. If the matrix in Eq. (13.1.5) is vertical, the right centering matrix in parentheses will be of the same order as the width of the data matrix. Its rank, however, will be one less than its order; otherwise, it would not be orthogonal to the unit vector. It can also be proved that it is of rank only one less by showing that it cannot be orthogonal to any other than the unit vector.

As already indicated, ipsative measures, whether derived from experimental or from computational procedures, cannot be normative in the sense

that one person can be compared with another with respect to a single variable. This is, of course, because some arbitrary constant has been added or subtracted from all scores of each subject measured. Since this is not in general the same constant for all subjects, the resulting measures are not comparable from one subject to another. This limitation of ipsative measures is often overlooked. Frequently, persons are erroneously compared with one another with respect to ipsative measures.

13.3 Basic Structure and the Problem of Origin

We have indicated in Section 13.1 of this chapter that in general the factor analytic results will vary for a given data matrix according to what is done about the problem of origin. The importance of this fact is not as widely recognized as it should be. We shall therefore examine the effect of row and column origin transformations on the basic structure characteristics of the data matrix.

In what follows, we shall not be concerned with whether the scale units are the same for each of the variables in the set. We shall assume for the time being that some sort of rational or natural scaling is available, or that the variables have been scaled in terms of standard deviation or equal variance units. The problem of scale will be considered in more detail in Chapter 15, even though the influence of scaling on basic structure is currently not well understood.

In the following sections we shall consider four cases. The first of these assumes that we have a left centered matrix in equal standard deviation units, and that we have already available the basic structure solution for the corresponding correlation or normative covariance matrix. On the basis of this solution we wish to determine the basic structure of the raw covariance matrix, i.e., the covariance matrix of the data matrix prior to left

centering.

The second case assumes that we have the basic structure factors of the covariance matrix prior to left centering, and we wish to find the basic structure matrices of the correlation matrix as functions of the basic structure factors of the raw covariance matrix.

The third case assumes that the deviation or normative data matrix has also been centered on the right. We shall investigate the basic structure of the minor product moment ipsative covariance matrix of this ipsative data matrix, as a function of the basic structure factors of the correlation or normative covariance matrix.

In the fourth case we have the basic structure of the covariance matrix obtained from the ipsative data matrix, and from this we wish to calculate the basic structure of the correlation or normative covariance matrix.

We shall now consider some relationships among the four types of covariance matrices.

13.4 Basic Structure of Raw Covariance Matrix from Correlation Matrix

13.4.1 Computational Equations

13.4.1a Definition of Notation

$\underline{X}$ is the raw score matrix.

$\underline{x}$ is the deviation score matrix.

$\underline{M}'$ is the vector of means.

$\underline{C}$ is the normative covariance matrix.

$\underline{G}$ is the raw covariance matrix.

$\underline{Q}$ is the basic orthonormal of $\underline{C}$.

$\underline{H}$ is the basic orthonormal of $\underline{G}$.

$\underline{\delta}$ is the basic diagonal of $\underline{C}$.

$\underline{\phi}$ is the basic diagonal of $\underline{G}$.

13.4.1b The Equations

$$x = (1 - \frac{1\,1'}{N})\,X \qquad (13.4.1)$$

$$x = P \Delta Q' \qquad (13.4.2)$$

$$M' = \frac{1'\,X}{N} \qquad (13.4.3)$$

$$C = x'\,x = Q\,\delta\,Q' \qquad (13.4.4)$$

$$G = X'\,X = H\,\phi\,H' \qquad (13.4.5)$$

$$V = Q'\,M \qquad (13.4.6)$$

If L = 1

$$_1Y_L = V'\,V + \delta_1 \qquad (13.4.7)$$

$$_1y_L = \delta_1 \qquad (13.4.8)$$

If L > 1

$$_1Y_L = \delta_{L-1} \qquad (13.4.9)$$

$$_1y_L = \delta_L \qquad (13.4.10)$$

$$_k Z_L = \frac{_k Y_L + _k y_L}{2} \tag{13.4.11}$$

$$_k F_L = \sum_{i=1}^{n} \frac{v_i^2}{\delta_i - _k Z_L} + 1 \tag{13.4.12}$$

If $\left| _k F_L \right| - P < 0$

$$\phi_L = _k Z_L \tag{13.4.13}$$

If $_k F_L < 0$

$$_{k+1} y_L = _k Z_L \tag{13.4.14}$$

$$_{k+1} Y_L = _k Y_L \tag{13.4.15}$$

If $_k F_L > 0$

$$_{k+1} Y_L = _k Z_L \tag{13.4.16}$$

$$_{k+1} y_L = _k y_L \tag{13.4.17}$$

$$f_{.i} = (\delta - \phi_i I)^{-1} v \tag{13.4.18}$$

$$H = Q(f D_{f'f}^{-\frac{1}{2}}) \tag{13.4.19}$$

$$a_G = H \phi^{\frac{1}{2}} \tag{13.4.20}$$

13.4.2 Computational Instructions. We begin with a score matrix $\underline{X}$, and for the sake of simplicity we assume that the measures all have unit standard deviations. We then apply a left centering matrix as in Eq. (13.4.1). This, as we know, simply subtracts the mean of each of the measures from every element in that vector.

Next we indicate the basic structure of this standard score matrix $\underline{x}$, as in Eq. (13.4.2). The right hand side is the product, from left to right, of the left orthonormal by the basic diagonal by the right orthonormal.

The means of the columns of the $\underline{X}$ matrix are indicated in Eq. (13.4.3). This is simply a row vector of the means of the variables.

The covariance minor product moment of the matrix in Eq. (13.4.2) is given by Eq. (13.4.4). Obviously, the left basic orthonormal disappears and the diagonal $\underline{\delta}$ is the square of $\underline{\Delta}$ in Eq. (13.4.2).

Eq. (13.4.5) is not a computational equation, but is given to indicate the minor product moment of $\underline{X}$ in terms of its basic structure. The basic orthonormals are $\underline{H}$ and $\underline{H}'$, respectively, and the basic diagonal is $\underline{\phi}$. We assume that the $\underline{Q}$ and the $\underline{\delta}$ matrices in Eq. (13.4.4) have been calculated according to one of the basic structure methods of factor analysis indicated in previous chapters. The problem now is to find the basic structure factors of the $\underline{G}$ matrix as a function of the known basic structure factors.

First, we calculate the $\underline{V}$ vector as in Eq. (13.4.6). This vector is the transpose of the $\underline{Q}$ matrix calculated in Eq. (13.4.4), postmultiplied by the vector of means calculated in Eq. (13.4.3).

To solve for the $\underline{\phi_L}$ elements of $\underline{\phi}$ we proceed as follows.

If $\underline{L} = 1$ we calculate a scalar $\underline{{}_1Y_1}$ as the sum of the minor product moment of $\underline{V}$ in Eq. (13.4.6) and the first element of $\underline{\delta}$, viz., $\underline{\delta_1}$, as shown in Eq. (13.4.7).

We then set a scalar $\underline{{}_1y_L}$ equal to $\underline{\delta_1}$ as indicated in Eq. (13.4.8).

If $\underline{L}$ is greater than 1 we set $\underline{{}_1Y_L}$ equal to $\underline{\delta_{L-1}}$, as in Eq. (13.4.9), and we set $\underline{{}_1y_L}$ equal to $\underline{\delta_L}$, as shown in Eq. (13.4.10).

In any case, as shown in Eq. (13.4.11), we set a scalar ${}_{k}Z_{L}$ equal to the average of ${}_{k}Y_{L}$ and ${}_{k}y_{L}$ where the latter two are determined iteratively, as we shall show shortly, for k going from 1 to some prespecified iteration limit.

We substitute ${}_{k}Z_{L}$ in Eq. (13.4.12).

If ${}_{k}F_{L}$ is sufficiently close to zero, as indicated by some tolerance limit P, we take ${}_{k}Z_{L}$ as the value of ρ_{L}, as shown in Eq. (13.4.13).

If ${}_{k}F_{L}$ is negative, we set a new y equal to ${}_{k}Z_{L}$ and the new Y will be the same as the previous one, as shown in Eqs. (13.4.14) and (13.4.15).

If ${}_{k}F_{L}$ is positive, we set the new Y equal to the current Z and the new y is the same as the previous one, as shown in Eqs. (13.4.16) and (13.4.17).

Once all of the ρ's have been solved for, we solve for vectors which are proportional to the vectors of H given in Eq. (13.4.5) by means of Eq.(13.4.18). This gives the solution for the ith column of an f matrix. As will be seen on the right of this equation, the V vector defined by Eq. (13.4.6) is premultiplied by the inverse of a diagonal matrix. This diagonal matrix for the ith column of the f matrix is obtained by subtracting from the δ matrix of Eq. (13.4.4) the ρ_{i} already solved for.

The f vector solved for in Eq. (13.4.18) is then normalized, as indicated by the product in parentheses on the right of Eq. (13.4.19). In addition, the normalized f matrix must also be premultiplied by the Q matrix solved for in Eq. (13.4.4) to yield the H matrix of Eq. (13.4.19).

To get the principal axis factor loading matrix for the G matrix in Eq. (13.4.5), the H matrix calculated in Eq. (13.4.19) must be postmultiplied by the square root of the diagonal matrix of ρ values. This is indicated in Eq. (13.4.20).

13.4.3 Numerical Example

In this example we shall use the same correlation matrix as in previous chapters. The vector of means used to generate the raw covariance matrix is given in Table 13.4.1.

We shall present the three covariance matrices - raw, normative, and ipsative, as well as their basic structures. The ipsatizing vector in the last two examples was taken as the unit vector divided by $\sqrt{n}$.

Table 13.4.2 gives the raw covariance matrix $\underline{G}$ which was obtained by adding the major product moment of the vector in Table 13.4.1 to the correlation matrix.

The first row of Table 13.4.3 gives the basic diagonal elements, $\underline{d}$, of the raw covariance matrix. The body of the table gives the basic orthonormal $\underline{H}$ of the raw covariance matrix in Table 13.4.2. These were calculated from $\underline{G}$ by a method in Chapter 9, as were also Tables 13.4.5 and 13.4.7.

Table 13.4.4 repeats for convenient reference the correlation matrix $\underline{R}$.

The first row of Table 13.4.5 gives the basic diagonal elements $\underline{\delta}$ of the correlation matrix $\underline{R}$. The body of the table gives the basic orthonormal matrix $\underline{Q}$ of $\underline{R}$.

Table 13.4.6 gives the ipsative covariance matrix calculated from the correlation matrix by the equation $\rho = (I - \frac{1\,1'}{n})\,R\,(I - \frac{1\,1'}{n})$.

The first row of Table 13.4.7 gives the basic diagonal elements of the ipsative correlation matrix $\underline{\rho}$. The body of the table gives the basic orthonormal matrix $\underline{q}$ of $\underline{\rho}$.

The first row of Table 13.4.8 gives the basic diagonal elements, $\underline{d}$, of the raw covariance matrix as calculated by Eqs. (13.4.7) through (13.4.17).

The body of the table gives the basic orthonormal $\underline{\underline{H}}'$ as calculated from Eqs. (13.4.18) and (13.4.19). Note that this matrix is the transpose of the one in Table 13.4.3. The values in the two tables agree within limits of decimal accuracy.

Table 13.4.1 - The Vector of Means of the Data Matrix

-0.4	-0.3	-0.2	-0.1	-0.0	0.1	0.2	0.3	0.4

Table 13.4.2 - The Raw Covariance Matrix C

1.16000	0.94900	0.84800	0.14800	0.03300	0.06800	0.21800	0.18900	0.19100
0.94900	1.09000	0.83500	0.14500	0.06100	0.09500	0.26300	0.25700	0.24900
0.84800	0.83500	1.04000	0.29200	0.20500	0.21800	0.25600	0.21100	0.30500
0.14800	0.14500	0.29200	1.01000	0.63600	0.61600	0.22900	0.15300	0.32900
0.03300	0.06100	0.20500	0.63600	1.00000	0.70900	0.13800	0.09100	0.25400
0.06800	0.09500	0.21800	0.61600	0.70900	1.01000	0.21000	0.13300	0.33100
0.21800	0.26300	0.25600	0.22900	0.13800	0.21000	1.04000	0.71400	0.60700
0.18900	0.25700	0.21100	0.15300	0.09100	0.13300	0.71400	1.09000	0.66100
0.19100	0.24900	0.30500	0.32900	0.25400	0.33100	0.60700	0.66100	1.16000

Table 13.4.3 - Basic Diagonals γ, and Basic Orthonormal H of the Raw Covariance Matrix

3.75072	2.18721	1.72066	0.50400	0.40746	0.34773	0.28544	0.22741	0.16937

0.37762	0.45666	0.17015	0.04074	0.00506	0.00084	0.11851	-0.45801	-0.62802
0.38813	0.41650	0.12300	0.02755	0.10607	0.06037	0.05831	-0.24569	0.76246
0.40369	0.29841	0.19922	-0.08977	-0.03480	-0.04380	-0.21507	0.79846	-0.11536
0.28681	-0.34966	0.29652	0.09157	-0.80422	0.10769	0.18423	-0.04775	0.06773
0.23789	-0.41259	0.35441	0.10045	0.29603	0.24950	-0.66799	-0.20104	-0.03115
0.26485	-0.40862	0.30373	0.02755	0.48419	-0.24399	0.60225	0.11329	-0.01920
0.32623	-0.12108	-0.44122	0.49806	-0.06532	-0.61914	-0.21504	-0.04648	0.01284
0.31173	-0.10407	-0.52983	0.26521	0.11557	0.68133	0.20372	0.13054	-0.06949
0.36221	-0.20973	-0.36750	-0.80740	-0.03244	-0.11623	-0.07552	-0.13262	-0.00460

Table 13.4.4 - The Normative Covariance or Correlation Matrix R

1.00000	0.82900	0.76800	0.10800	0.03300	0.10800	0.29800	0.30900	0.35100
0.82900	1.00000	0.77500	0.11500	0.06100	0.12500	0.32300	0.34700	0.36900
0.76800	0.77500	1.00000	0.27200	0.20500	0.23800	0.29600	0.27100	0.38500
0.10800	0.11500	0.27200	1.00000	0.63600	0.62600	0.24900	0.18300	0.36900
0.03300	0.06100	0.20500	0.63600	1.00000	0.70900	0.13800	0.09100	0.25400
0.10800	0.12500	0.23800	0.62600	0.70900	1.00000	0.19000	0.10300	0.29100
0.29800	0.32300	0.29600	0.24900	0.13800	0.19000	1.00000	0.65400	0.52700
0.30900	0.34700	0.27100	0.18300	0.09100	0.10300	0.65400	1.00000	0.54100
0.35100	0.36900	0.38500	0.36900	0.25400	0.29100	0.52700	0.54100	1.00000

Table 13.4.5 - Basic Diagonals δ, and Basic Orthonormal Q of the Correlation Matrix

3.74907	2.04953	1.33079	0.47442	0.38261	0.34740	0.28533	0.21215	0.16870

0.37009	0.34414	-0.30364	0.04349	0.01542	0.00399	-0.12630	-0.44249	-0.66221
0.38211	0.33393	-0.27878	0.08123	0.07647	-0.05713	-0.06068	-0.31066	0.74085
0.39915	0.20657	-0.35222	-0.01294	-0.14028	0.02245	0.23294	0.77085	-0.07157
0.28690	-0.45323	-0.05853	-0.14862	-0.78182	-0.15550	-0.17400	-0.13710	0.04754
0.23932	-0.51976	-0.15671	0.16644	0.29930	-0.22904	0.66174	-0.19627	-0.04304
0.26733	-0.48451	-0.16335	0.18680	0.40810	0.26448	-0.60582	0.16939	-0.00176
0.33081	0.05559	0.50930	0.41816	-0.18573	0.60692	0.21552	-0.06236	0.00760
0.31760	0.11592	0.53843	0.23892	0.11339	-0.67780	-0.20136	0.15045	-0.05702
0.36926	-0.02350	0.31947	-0.82141	0.24394	0.14790	0.06259	-0.03072	0.00889

Table 13.4.6 - The Ipsative Covariance Matrix $\underline{\rho}$

0.56570	0.37915	0.28859	-0.29896	-0.32607	-0.28030	-0.12196	-0.09141	-0.11474
0.37915	0.53459	0.28004	-0.30752	-0.31363	-0.27885	-0.11252	-0.06896	-0.11230
0.28859	0.28004	0.47548	-0.18007	-0.19919	-0.19541	-0.16907	-0.17452	-0.12585
-0.29896	-0.30752	-0.18007	0.62037	0.30426	0.26504	-0.14363	-0.19007	-0.06941
-0.32607	-0.31363	-0.19919	0.30426	0.71615	0.39593	-0.20674	-0.23419	-0.13652
-0.28030	-0.27885	-0.19541	0.26504	0.39593	0.65770	-0.18396	-0.25141	-0.12874
-0.12196	-0.11252	-0.16907	-0.14363	-0.20674	-0.18396	0.59437	0.26793	0.07559
-0.09141	-0.06896	-0.17452	-0.19007	-0.23419	-0.25141	0.26793	0.63348	0.10915
-0.11474	-0.11230	-0.12585	-0.06941	-0.13652	-0.12874	0.07559	0.10915	0.50281

Table 13.4.7 - Basic Diagonals $\underline{d}$, and Basic Orthonormal of the Ipsative Covariance Matrix $\underline{\rho}$.

2.08395	1.33155	0.48297	0.38347	0.34929	0.28534	0.21507	0.16904	-0.00000

-0.39228	0.30627	-0.04279	-0.01569	0.00150	-0.12549	0.44869	-0.64965	0.33334
-0.38541	0.28215	-0.07309	-0.07283	-0.04748	-0.06101	0.28349	0.75083	0.33334
-0.26670	0.35823	0.04678	0.15209	0.03581	0.23107	-0.77588	-0.08622	0.33334
0.40533	0.06788	0.18405	0.77628	-0.18170	-0.17375	0.13256	0.05212	0.33333
0.48304	0.16394	-0.16644	-0.29888	-0.20860	0.66065	0.17497	-0.03241	0.33334
0.44111	0.17168	-0.16513	-0.38436	0.29794	-0.60680	-0.18124	-0.00148	0.33334
-0.09835	-0.50431	-0.37720	0.22433	0.61827	0.21600	0.05703	0.01119	0.33333
-0.15311	-0.53508	-0.24311	-0.12428	-0.65532	-0.20398	-0.17530	-0.05301	0.33328
-0.03365	-0.31090	0.83691	-0.25666	0.13948	0.06331	0.03572	0.00862	0.33333

Table 13.4.8 - Basic Diagonal Elements $\underline{d}$, and Basic Orthonormal $\underline{H}'$ as Determined from Basic Structure of $\underline{R}$

3.75072	2.18720	1.72065	0.50401	0.40746	0.34773	0.28544	0.22742	0.16938

0.3776	0.3882	0.4037	0.2868	0.2380	0.2648	0.3262	0.3117	0.3622
0.4566	0.4165	0.2984	-0.3497	-0.4127	-0.4086	-0.1211	-0.1041	-0.2097
0.1701	0.1230	0.1992	0.2965	0.3544	0.3037	-0.4412	-0.5298	-0.3675
0.0408	0.0276	-0.0897	0.0915	0.1005	0.0275	0.4980	0.2652	-0.8074
-0.0050	-0.1061	0.0348	0.8042	-0.2960	-0.4842	0.0653	-0.1156	0.0325
0.0008	0.0604	-0.0438	0.1077	0.2495	-0.2440	-0.6192	0.6813	-0.1162
-0.1185	-0.0584	0.2151	-0.1843	0.6679	-0.6023	0.2150	-0.2037	0.0755
0.4580	0.2457	-0.7984	0.0477	0.2011	-0.1133	0.0465	-0.1306	0.1326
0.6281	-0.7624	0.1153	-0.0677	0.0312	0.0192	-0.0129	0.0695	0.0046

13.5 The Normative Covariance Basic Structure from the Raw Covariance Basic Structure

13.5.1 Computational Equations

13.5.1a Definition of Notation

The notation is the same as in the previous section.

13.5.1b The Equations

$$V = H' M \tag{13.5.1}$$

$$_1Y_n = \phi_n \tag{13.5.2}$$

$$_1y_n = 0 \tag{13.5.3}$$

$$_1Y_L = \phi_L \tag{13.5.4}$$

$$_1y_L = \phi_{L+1} \tag{13.5.5}$$

$$_kZ_L = \frac{_kY_L + {}_ky_L}{2} \tag{13.5.6}$$

$$_kF_L = \sum_{i=1}^{n} \frac{V_i^2}{\phi_i - {}_kZ_L} - 1 \tag{13.5.7}$$

$$W_{.L} = H(\phi - \delta_L I)^{-1} V \tag{13.5.8}$$

$$Q_{.L} = \frac{W_{.L}}{\sqrt{W'_{.L} W_{.L}}} \tag{13.5.9}$$

13.5.2 Computational Instructions

First we calculate the vector $\underline{V}$ of Eq. (13.5.1). This is the vector of means premultiplied by the right orthonormal of the raw covariance matrix, $\underline{Q}$.

Eqs. (13.5.2) and (13.5.3) give the limits of the normative basic diagonal, $\underline{\delta_n}$. Eqs. (13.5.4) and (13.5.5) give the limits of the normative basic

diagonals for δ_L where L is less than n.

Eq. (13.5.6) gives the kth approximation to the Lth basic diagonal, δ_L.

To solve for the Lth basic diagonal of ${}_1\delta_L$, we use Eq. (13.5.7) iteratively, as we used Eq. (13.4.12) to solve for ϕ_L. We note, however, that the last term on the right in Eq. (13.4.12) is +1, whereas it is -1 for Eq. (13.5.7).

Having solved for the δ's, we substitute these in Eq. (13.5.8) to solve for a vector proportional to the Lth vector of Q, the basic orthonormal of the correlation or normative covariance matrix.

Eq. (13.5.9) shows the normalization of the $W_{.L}$ vector of Eq. (13.5.8) to give the Lth vector of Q.

13.5.3 Numerical Example

The first row of Table 13.5.1 gives the basic diagonal elements δ of the correlation matrix R as calculated by Eqs. (13.5.1) through (13.5.7). The body of the table gives the basic orthonormal Q' of R as calculated by Eqs. (13.5.8) and (13.5.9). Note that this matrix is the transpose of the one in Table 13.4.5. The values in the two tables agree within limits of decimal accuracy.

13.6 The Ipsative Covariance Basic Structure from the Normative Covariance Basic Structure

13.6.1 Computational Equations

13.6.1a Definition of Notation

R is the normative covariance matrix.

Q is the basic orthonormal of R.

$\underline{\delta}$ is the basic diagonal of $\underline{R}$.

$\underline{\rho}$ is the ipsative covariance matrix.

$\underline{q}$ is the basic orthonormal of $\underline{\rho}$.

$\underline{d}$ is the basic diagonal of $\underline{\rho}$.

$\underline{V}$ is a normal vector whose order is the same as $\underline{R}$.

The relation between $\underline{\rho}$ and $\underline{R}$ is given by $\rho = (I - V V') R (I - V V')$.

13.6.1b The Equations

$$U = Q' V \tag{13.6.1}$$

$$W = \delta^{\frac{1}{2}} U \tag{13.6.2}$$

$$_1Y_n = \delta_n \tag{13.6.3}$$

$$_1y_n = 0 \tag{13.6.4}$$

$$_1Y_L = \delta_L \tag{13.6.5}$$

$$_1y_L = \delta_{L+1} \tag{13.6.6}$$

$$_kZ_L = \frac{_kY_L + {_ky_L}}{2} \tag{13.6.7}$$

$$_kF_L = \sum_{i=1}^{n} \frac{W_L^2}{\delta_i - {_kZ_L}} - 1 \tag{13.6.8}$$

$$q_{.L} = Q (\delta - d_L I)^{-1} U g \tag{13.6.9}$$

Table 13.5.1 - Basic Diagonals $\underline{\delta}$, and Basic Orthonormal $\underline{Q}'$ as Determined from $\underline{G}$

3.74907	2.04952	1.33077	0.47442	0.38262	0.34740	0.28533	0.21215	0.16871
-0.3701	-0.3821	-0.3991	-0.2869	-0.2392	-0.2674	-0.3308	-0.3176	-0.3693
-0.3442	-0.3339	-0.2066	0.4532	0.5197	0.4846	-0.0556	-0.1159	0.0235
-0.3037	-0.2787	-0.3522	-0.0586	-0.1567	-0.1633	0.5093	0.5384	0.3195
-0.0435	-0.0812	0.0129	0.1486	-0.1664	-0.1868	-0.4182	-0.2389	0.8214
0.0154	0.0765	-0.1403	-0.7818	0.2993	0.4081	-0.1858	0.1134	0.2440
0.0039	-0.0571	0.0224	-0.1555	-0.2291	0.2644	0.6069	-0.6778	0.1479
0.1263	0.0606	-0.2329	0.1739	-0.6618	0.6058	-0.2156	0.2013	-0.0626
-0.4425	-0.3106	0.7709	-0.1371	-0.1962	0.1693	-0.0623	0.1504	-0.0307
-0.6622	0.7409	-0.0715	0.0476	-0.0431	-0.0017	0.0075	-0.0570	0.0089

13.6.2 Computational Instructions

Eq. (13.6.1) defines a vector $\underline{U}$ which is the product of the right orthonormal matrix of $\underline{R}$ postmultiplied by the ipsatizing vector $\underline{V}$.

Eq. (13.6.2) defines a vector $\underline{W}$ which is the product of the $\underline{U}$ vector in Eq. (13.6.1) premultiplied by the square root of the basic diagonal of $\underline{R}$.

Eqs. (13.6.3) through (13.6.6) give the first approximations to the limits of the basic diagonals of $\underline{\rho}$, the ipsatized covariance matrix.

Eq. (13.6.7) gives the $\underline{k}$th approximation to the $\underline{L}$th basic diagonal of $\underline{\rho}$.

The iteration procedure for getting successively smaller bounds for the $\underline{d}$'s is the same as in the two previous methods, except that now Eq. (13.6.8) is the iteration equation. It is of the same form as in the previous two methods. In this case the "1" is subtracted on the right as in Eq. (13.5.7).

Eq. (13.6.9) gives the calculations for the $\underline{q}_{.L}$ vectors of the basic orthonormal $\underline{q}$ of $\underline{\rho}$. The factor $\underline{g}$ on the extreme right is a normalizing scalar.

13.6.3 Numerical Example

The first row of Table 13.6.1 gives the basic diagonal elements $\underline{d}$ of the ipsative matrix $\underline{\rho}$ as calculated from Eqs. (13.6.1) through (13.6.8). The body of the table gives the basic orthonormal $\underline{q}'$ of $\underline{\rho}$ as calculated from Eq. (13.6.9). Note that this matrix is the transpose of the one in Table 13.4.7. The values in both tables agree within limits of decimal accuracy, except the last line of Table 13.6.1. This discrepancy is due to the error in the last basic diagonal which should be 0 instead of .00002.

Table 13.6.1 - Basic Diagonals $\underline{d}$, and Basic Orthonormal $\underline{q}'$ as Determined from Basic Structure of $\underline{R}$

2.08395	1.33155	0.48296	0.38347	0.34928	0.28534	0.21506	0.16903	0.00002
0.3923	0.3854	0.2667	-0.4053	-0.4830	-0.4411	0.0983	0.1531	0.0336
0.3062	0.2821	0.3583	0.0679	0.1640	0.1717	-0.5043	-0.5351	-0.3109
-0.0428	-0.0731	0.0468	0.1841	-0.1664	-0.1651	-0.3772	-0.2431	0.8369
-0.0157	-0.0728	0.1521	0.7763	-0.2988	-0.3844	0.2243	-0.1243	-0.2567
0.0015	-0.0475	0.0359	-0.1817	-0.2085	0.2979	0.6183	-0.6553	0.1395
0.1255	0.0611	-0.2310	0.1738	-0.6606	0.6069	-0.2160	0.2041	-0.0633
-0.4487	-0.2835	0.7758	-0.1325	-0.1750	0.1813	-0.0570	0.1754	-0.0357
-0.6496	0.7508	-0.0862	0.0521	-0.0324	-0.0015	0.0112	-0.0530	0.0086
0.5171	0.2582	-0.0732	0.2448	0.5206	0.2746	0.2746	0.4190	0.0571

13.7 The Normative Covariance Basic Structure from the Ipsative Basic Structure

13.7.1 Computational Equations

13.7.1a Definition of Notation

The notation is the same as in Section 13.6.1.

13.7.1b The Equations

$$U = R\ V \tag{13.7.1}$$

$$W = q'\ U \tag{13.7.2}$$

$$\alpha = V'\ U \tag{13.7.3}$$

$${}_1Y_1 = \frac{d_1 + n}{2} \tag{13.7.4}$$

$${}_1y_1 = d_1 \tag{13.7.5}$$

$${}_1Y_L = d_{L-1} \tag{13.7.6}$$

$${}_1y_L = d_L \tag{13.7.7}$$

$${}_kZ_L = \frac{{}_kY_L + {}_ky_L}{2} \tag{13.7.8}$$

$${}_kF_L = \sum_{L=1}^{n-1} \frac{W_L^2}{d_L - {}_kZ_L} + {}_kZ_L - \alpha \tag{13.7.9}$$

$$Q_{.L} = (V - q\ (d - \delta_L\ I)^{-1}\ W)g_L \tag{13.7.10}$$

13.7.2 Computational Instructions

Eq. (13.7.1) gives a vector $\underline{U}$ as the product of the normative covariance matrix postmultiplied by the ipsatizing vector $\underline{V}$. If one has only the $\underline{\rho}$ matrix to begin with, as in the case of ipsative personality measures, then

one may be able to hypothesize a $\underline{U}$ vector, for the computational procedure requires only the $\underline{U}$ vector and not the $\underline{R}$ matrix as such.

Eq. (13.7.2) gives a $\underline{W}$ vector as the product of the right orthonormal $\underline{q}'$ of $\underline{\rho}$ and the $\underline{U}$ vector of Eq. (13.7.1).

The next step is the calculation of the scalar $\underline{\alpha}$ in Eq. (13.7.3). This is the minor product of the $\underline{U}$ and $\underline{V}$ vectors.

The outer limits of the first basic diagonal of $\underline{R}$, viz. $\underline{\delta_1}$, are given by Eqs. (13.7.4) and (13.7.5). The $\underline{{}_1Y_1}$ value assumes that the normative matrix is actually a correlation matrix so that its trace is $\underline{n}$, the order of the matrix. It is well known that this trace is the sum of the basic diagonals, hence $\underline{\delta_1}$ must be less than $\underline{n}$.

Eqs. (13.7.6) and (13.7.7) give the outer limits of the remaining $\underline{\delta_L}$ values.

As in the solutions of the previous sections, the $\underline{k}$th approximation to $\underline{\delta_L}$ is given by $\underline{{}_kZ_L}$ in Eq. (13.7.8).

Eq. (13.7.9) gives the iteration equation for the $\underline{\delta_L}$ values. The same procedure is used for narrowing the limits of the $\underline{\delta_L}$'s as in the previous sections.

It is to be noted, however, that the summation goes only to $\underline{n}$-1. This equation also differs from the corresponding equation of previous sections in that the right hand side includes the $\underline{Z}$ and $\underline{\alpha}$ terms instead of "1".

Eq. (13.7.10) shows the calculations for the $\underline{Q_{.L}}$ vectors of the basic orthonormal of $\underline{R}$. The $\underline{g_L}$ on the extreme right is a normalizing scalar.

13.7.3 Numerical Example

The first row of Table 13.7.1 gives the basic diagonal elements $\underline{\delta}$ of the correlation matrix $\underline{R}$, as calculated from Eqs. (13.7.1) through (13.7.9).

The body of the table gives the basic orthonormal $\underline{Q}'$ of $\underline{R}$, as calculated from Eq. (13.7.10.). Note that this matrix is the transpose of the one in Table 13.4.5. The values in the two tables agree within limits of decimal accuracy.

13.8 Mathematical Proofs

13.8.1 Basic Structure and Left Centering

Given the matrix

$$x = \left(I - \frac{1\,1'}{N}\right) X \tag{13.8.1}$$

and

$$x = P \Delta Q' \tag{13.8.2}$$

$$M' = \frac{1' X}{N} \tag{13.8.3}$$

From Eqs. (13.8.1) and (13.8.3)

$$x' x + M M' = X' X \tag{13.8.4}$$

Let

$$C = x' x \tag{13.8.5}$$

$$G = X' X \tag{13.8.6}$$

From Eqs. (13.8.4), (13.8.5), and (13.8.6)

$$C + M M' = G \tag{13.8.7}$$

From Eqs. (13.8.2), (13.8.3), and (13.8.7)

$$Q \Delta^2 Q' + M M' = G \tag{13.8.8}$$

Let us now find the roots of $\underline{G}$.

Assume $\underline{Q}$ is vertical. Let $\underline{Q} = \underline{Q_i}$ and $\underline{Q_j}$ be a complement of $\underline{Q_i}$. In particular,

Table 13.7.1 - Basic Diagonals $\underline{\delta}$, and Basic Orthonormal $\underline{Q}'$ as Determined from Basic Structure of $\underline{\rho}$ and from $\underline{R}\ \underline{V}$

3.74906	2.04952	1.33080	0.47443	0.38261	0.34741	0.28533	0.21215	0.16870
0.3701	0.3821	0.3991	0.2069	0.2392	0.2673	0.3308	0.3177	0.3693
-0.3442	-0.3340	-0.2066	0.4532	0.5197	0.4843	-0.0556	-0.1159	0.0235
-0.3037	-0.2788	-0.3521	-0.0585	-0.1567	-0.1633	0.5093	0.5385	0.3195
0.0435	0.0812	-0.0129	-0.1486	0.1665	0.1868	0.4181	0.2390	-0.8214
0.0154	0.0765	-0.1403	-0.7818	0.2993	0.4081	-0.1858	0.1134	0.2439
-0.0039	0.0572	-0.0224	0.1555	0.2291	-0.2644	-0.6069	0.6778	-0.1479
-0.1263	-0.0606	0.2330	-0.1740	0.6618	-0.6058	0.2156	-0.2014	0.0626
0.4425	0.3106	-0.7709	0.1371	0.1963	-0.1694	0.0624	-0.1504	0.0307
0.6622	-0.7408	0.0716	-0.0476	0.0431	0.0018	-0.0076	0.0570	-0.0089

we may have

$$Q_j\, Q_j' = I - Q_i\, Q_i' \qquad (13.8.9)$$

where $\underline{Q_j}$ is the partial triangular factor of the right of Eq. (13.8.9).

From Eq. (13.8.8) we may write

$$\left[(Q_i\ ,\ Q_j) \begin{bmatrix} \Delta^2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} Q_i' \\ Q_j' \end{bmatrix} + M\,M' - \phi_i\, I \right] H_{.i} = 0 \qquad (13.8.10)$$

where $\underline{H_{.i}}$ is a basic vector of $\underline{G}$.

From Eq. (13.8.10)

$$\left[\begin{bmatrix} \Delta^2 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} Q_i' \\ Q_j' \end{bmatrix} M\,M' \,(Q_i\ ,\ Q_j) - \begin{bmatrix} \phi_i\, I & 0 \\ 0 & \phi_I\, I \end{bmatrix} \right] \begin{bmatrix} Q_i' \\ Q_j' \end{bmatrix} H_{.i} = 0$$

(13.8.11)

Let

$$Q_i'\, H_{.i} = V_{.i} \qquad (13.8.12)$$

$$Q_j'\, H_{.i} = V_{.j} \qquad (13.8.13)$$

$$Q_i'\, M = V_i \qquad (13.8.14)$$

$$Q_j'\, M = V_j \qquad (13.8.15)$$

Eqs. (13.8.12) through (13.8.15) in (13.8.11) give

$$\left[\begin{bmatrix} \Delta^2 - \phi I & 0 \\ 0 & -\phi I \end{bmatrix} + \begin{bmatrix} V_i \\ V_j \end{bmatrix} (V_i', V_j')\right] \begin{bmatrix} V_{.i} \\ V_{.j} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (13.8.16)$$

From Eq. (13.8.16)

$$\begin{bmatrix} V_{.i} \\ V_{.j} \end{bmatrix} + \begin{bmatrix} (\Delta^2 - \phi I)^{-1} & 0 \\ 0 & -I\frac{1}{\phi} \end{bmatrix} \begin{bmatrix} V_i \\ V_j \end{bmatrix} (V_i' V_{.i} + V_j' V_{.j}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

(13.8.17)

From Eq. (13.8.17)

$$1 + V_i' (\Delta^2 - \phi I)^{-1} V_i + \frac{V_j' V_j}{\phi} = 0 \quad (13.8.18)$$

In particular, if $\underline{c}$ is basic, then Eq. (13.8.18) becomes

$$1 + V' (\Delta^2 - \phi I)^{-1} V = 0 \quad (13.8.19)$$

where $\underline{i}$ is dropped from $\underline{V_i}$.

Or, in scalar notation,

$$1 + \frac{V_1^2}{\delta_1 - \phi} + \frac{V_2^2}{\delta_2 - \phi} + \dots + \frac{V_n^2}{\delta_n - \phi} = 0 \quad (13.8.20)$$

where the $\underline{V}$'s in the numerators of Eq. (13.8.20) are the elements of $\underline{V}$ in Eq. (13.8.19).

To solve for the $\underline{k}$th root $\underline{\phi_k}$ in Eq. (13.8.20) we consider

$$F = 1 + \frac{V_1^2}{\delta_1 - Z} + \frac{V_2^2}{\delta_2 - Z} + \dots + \frac{V_n^2}{\delta_n - Z} \quad (13.8.21)$$

We can now prove that a root of $\underline{F}$ lies between $\underline{\delta_{k+1}}$ and $\underline{\delta_k}$. As $\underline{Z}$ approaches $\underline{\delta_{k+1}}$ from above, $\underline{F} \to \underline{-\infty}$, and as it approaches $\underline{\delta_k}$ from below, $\underline{F} \to \underline{\infty}$.

Therefore a root of F must lie between δ_{k+1} and δ_k. To solve for any ϕ_k lying between δ_{k+1} and δ_k we may begin by letting

$$Y = \delta_k \qquad (13.8.22)$$

$$y = \delta_{k+1} \qquad (13.8.23)$$

$$Z = \frac{Y + y}{2} \qquad (13.8.24)$$

If now Z in Eq. (13.8.21) gives F positive, Z is too large and we take

$$Y = Z$$

If Z in Eq. (13.8.21) had given F negative, Z would have been too small and we would have taken

$$y = Z$$

and used Eq. (13.8.24) for a new Z. We continue in this manner until F in Eq. (13.8.21) is sufficiently close to 0.

We know, however, that

$$\phi_1 \geq \delta_1 \qquad (13.8.25)$$

Therefore we must find an upper bound to ϕ_1. This will in general not be greater than

$$Y = M' M + \delta_1 \qquad (13.8.26)$$

To solve for the $H_{.k}$ we have from Eq. (13.8.17)

$$V_{.k} = (\delta - \phi_k I)^{-1} V \qquad (13.8.27)$$

From Eq. (13.8.12)

$$U_{.k} = Q V_{.k} \tag{13.8.28}$$

and

$$H_{.k} = \frac{U_{.k}}{\sqrt{U'_{.k} U_{.k}}} \tag{13.8.29}$$

The proof for the left centered matrix from that of the uncentered matrix follows similar lines.

13.8.2 Proof of Basic Structure and Right Centering Ipsative from Normative Basic Structure Factors

Let $\underline{R}$ be the basic normative covariance or correlation matrix and $\underline{\rho}$ the ipsative covariance matrix so that

$$\rho = (I - V V') R (I - V V') \tag{13.8.30}$$

where

$$V' V = 1 \tag{13.8.31}$$

Let the basic structures of $\underline{R}$ and $\underline{\rho}$, respectively, be

$$R = Q \delta Q' \tag{13.8.32}$$

and

$$\rho = q d q' \tag{13.8.33}$$

From Eqs. (13.8.30), (13.8.32), and (13.8.33)

$$q d q' = (I - V V') Q \delta Q' (I - V V') \tag{13.8.34}$$

Because of Eq. (13.8.30)

$$V' q = 0 \tag{13.8.35}$$

From Eqs. (13.8.34) and (13.8.35)

$$q\, d - (I - V V') Q\, \delta\, Q' q = 0 \tag{13.8.36}$$

From Eq. (13.8.36)

$$Q' q\, d - (I - Q' V V' Q)\, \delta\, Q' q = 0 \tag{13.8.37}$$

From Eq. (13.8.37)

$$\delta\, Q' q - Q' q\, d - Q' V V' Q\, \delta\, Q' q = 0 \tag{13.8.38}$$

From Eq. (13.8.38)

$$Q' q_{.i} - (\delta - d_i I)^{-1} Q' V V' Q\, \delta\, Q' q_{.i} = 0 \tag{13.8.39}$$

From Eq. (13.8.39)

$$(V' Q\, \delta) Q' q_{.i} - (V' Q\, \delta)(\delta - d_i I)^{-1} Q' V V' Q\, \delta\, Q' q_{.i} = 0$$

or

$$V' Q\, \delta^{\frac{1}{2}} (\delta - d_i I)^{-1} \delta^{\frac{1}{2}} Q' V - 1 = 0 \tag{13.8.40}$$

From Eq. (13.8.40) we solve for the $\underline{d_i}$ as in the case of the raw-normative methods. Having solved for the basic diagonal, we can solve for the basic orthonormal of $\underline{\rho}$ from Eq. (13.8.39). We have

$$q_{.i} = Q (\delta - d_i I)^{-1} Q' V g \tag{13.8.41}$$

where g is a normalizing scalar.

Normative from ipsative basic structure factors.

From Eqs. (13.8.34) and (13.8.35) we have

$$d\, q'\, Q - q'\, Q\, \delta\, (I - Q'\, V\, V'\, Q) = 0 \qquad (13.8.42)$$

From Eqs. (13.8.32) and (13.8.42)

$$d\, q'\, Q_{.i} - q'\, Q_{.i}\, \delta_i + q'\, R\, V\, V'\, Q_{.i} = 0 \qquad (13.8.43)$$

From Eq. (13.8.43)

$$q'\, Q_{.i} + (d - \delta_i\, I)^{-1}\, q'\, R\, V\, V'\, Q_{.i} = 0 \qquad (13.8.44)$$

From Eq. (13.8.44)

$$q\, q'\, Q_{.i} + q\, (d - \delta_i\, I)^{-1}\, q'\, R\, V\, V'\, Q_{.i} = 0 \qquad (13.8.45)$$

It can be proved that

$$q\, q' = I - V\, V' \qquad (13.8.46)$$

From Eqs. (13.8.45) and (13.8.46)

$$Q_{.i} - V\, V'\, Q_{.i} + q\, (d - \delta_i\, I)^{-1}\, q'\, R\, V\, V'\, Q_{.i} = 0 \qquad (13.8.47)$$

Premultiplying Eq. (13.8.47) by $\underline{V'\, R}$ gives, because of Eq. (13.8.32),

$$V'\, Q_{.i}\, \delta_i - V'\, R\, V\, V'\, Q_{.i} + V'\, R\, q\, (d - \delta_i\, I)^{-1}\, q'\, R\, V\, V'\, Q_{.i} = 0$$

or

$$V'\, R\, q\, (d - \delta_i\, I)^{-1}\, q'\, R\, V + \delta_i - V'\, R\, V = 0 \qquad (13.8.48)$$

If we have the vector R V we can solve Eq. (13.8.48) for the δ_i's by methods analogous to the previous methods since we can readily show that

$$d_i \geq \delta_{i+1} \geq \delta_i \qquad (13.8.49)$$

We still need an upper bound for δ_1, however.

We can solve for the $Q_{.i}$ by rewriting Eq. (13.8.47) as

$$Q_{.i} = (V - q\,(d - \delta_i\, I)^{-1}\, q'\, R\, V)\, g_i \qquad (13.8.50)$$

where g_i is a normalizing scalar.

CHAPTER 14

CATEGORICAL VARIATIONS IN FACTOR ANALYSIS

In the previous chapters we have regarded data matrices as consisting essentially of measures for a given number of entities on each of a different number of attributes. This type of data matrix has been the one most commonly used in factor analytic studies. You recall, however, that in several chapters we indicated that in some cases the attributes might have a natural origin of measurement. For example, in the previous chapter we suggested that the entities might be persons practicing typewriting and the attributes might be successive time intervals.

14.1 Multicategory Sets

We may, however, consider a somewhat more general model. For example, suppose that we have a number of individuals for each of whom a number of physiological and psychological variables were measured on each of a number of successive days or time intervals. For such a set of data we have three, rather than two, categories to consider. In the conventional case we have entities and attributes only. In this more general case we have entities, attributes, and occasions. Let us now consider possible ways of studying data of this type.

14.1.1 The Attribute-Entity Sets. In the case of the three category set, we may consider a number of matrices of the conventional type consisting of attributes as columns and entities as rows. Each of these matrices would have the same attributes and entities for a number of different occasions. We may call these matrices slabs of the three category data matrix.

The problem of how to handle such a set of data by means of factor analytic technique is one which has not been thoroughly explored. Horst (1963) has recently discussed the general problem of multicategory sets of

data and proposed several different ways of analyzing such sets. Tucker (1963) also has considered the general problem of entities, attributes, and occasions in what he calls the three mode factor analysis model. He has presented an ingenious procedure for conceiving of a three category set of data in terms of what he calls a core matrix, which includes the categories of entities, attributes, and occasions. The method assumes lower orders for each of these categories than are represented in the data matrix and the problem then is to solve for this lower order three category matrix as a basis for reproducing the observed three category data matrix.

We may consider a simpler way of handling data of this type as a two category set. Here the analysis of the data would be amenable to the techniques which we have discussed in the previous chapters.

The first of these ways of considering the data is to regard each occasion for each attribute as a distinct attribute. Thus we would regard an attribute measured today for a group of entities as different from the same attribute measured tomorrow on the same entities. For example, the variable of typewriting speed on Monday for individual A and the variable of typewriting speed on Tuesday for the same individual would be regarded as two different variables. We would then consider a supermatrix in which the entity-attribute slabs would be strung out in such a way that, if we had 4 occasions and 10 attributes, the supermatrix would actually have 10 x 4, or 40, attributes. We would therefore have a 40-variable matrix. We may then consider a factor analysis of such a matrix along the lines outlined in the previous chapters. This could be solved for principal axis factor loadings for each variable or attribute on each occasion.

Another way of regarding the same set of data would be to consider each occasion as a different set of entities. This would mean that the person whose typewriting speed is recorded today will be considered a different person when his speed is recorded tomorrow, even though he has the same name and is identifiable as the same individual. If we look at the problem in this way, we could then string out the entity-attribute slabs for the various occasions so that, if we had, say, 20 persons and 4 occasions, we would actually now have 80 different persons since we regard each individual as a different person on each of the 4 different occasions. This would give us a supermatrix, or a column type 3 supervector, in which each of the matric elements is an entity-attribute matrix for a specific occasion. Here we would have 80 entities and 10 attributes. We can now, on a matrix of this type, do a factor analysis according to procedures described in previous chapters. Such an analysis could then yield a set of factor scores for each person on each of the 4 different occasions. It would also yield a set of factor loadings for each of the 10 variables.

It will be seen that the first and the second ways of setting up the matrices of data in the form of type 3 supervectors yield essentially different results. In the first case we get a factor score matrix for the 20 entities and we get factor loading matrices for the 10 attributes on each of 4 different occasions. In the second case we get a factor score matrix for the entities on each of the 4 different occasions and a factor loading matrix for the 10 attributes. In the one case we regard the occasions as different variables, and in the second we regard them as different entities.

14.1.2 The Attribute-Occasion Set. In the previous example we have considered the decomposition of the data cube, as it were, into occasion

slabs such that each slab had entities for rows and attributes for columns. We may now consider a different decomposition of this three dimensional matrix such that each slab represents a person. Each person matrix may be regarded as consisting of occasion rows and attribute columns. We would then have, using the previous example, twenty 4 x 10 matrices.

We now have two obvious alternatives of treating these slabs of data. In the first of these, we could string out the slabs into a row type 2 supervector so that each person is regarded as a different attribute. Therefore we would have a supermatrix of 4 occasions and 10 x 20, or 200, attributes. On such a matrix one could then perform a factor analysis.

On the other hand, we may string the matrices in the other direction, so that each person is considered as a different occasion. We would therefore have a matrix with 80 rows and 10 columns. We recognize at once that this arrangement of the data is the same as the second way of arranging it in the previous method, except that there has been an interchange of rows. In both methods the columns are the 10 attributes, but the entities for a single occasion are grouped together in rows. In the second case we have the same 10 attributes, but the occasion rows are grouped for a specific entity. We see, therefore, that we actually have, so far, three different ways of arranging the data into a row by column data matrix which can be factor analyzed by available methods.

14.1.3 The Entity-Occasion Pair or Set. Let us now see what happens if we take the third remaining possibility of decomposing the data. In the first case we took slabs such that each slab was a different occasion. In the second case we took slabs such that each slab was a different person or entity, and in the third case we take slabs from the cube so that each slab

is a different attribute.

In this latter case, each slab may be regarded as having entities for rows and occasions for columns. Suppose we string these slabs out into a row supervector so that the 20 entities constitute the rows, and the columns are sets of occasions for each of the successive attributes. We recognize at once that this is the first arrangement considered in the attribute-entity slabs, except that occasions are grouped by attributes, while in the first case attributes were grouped by successive occasions.

Obviously then, since a data matrix of this type has simply undergone a right hand permutation of the former type of data matrix, the factor analysis results would be the same except for the permutation of columns in the data matrix and the corresponding permutation of rows in the factor loading matrix.

Let us see now what happens if we string these entity-occasion slabs in a vertical manner so that in the column supervector we have occasions for columns, and the rows consist of entities grouped by successive attributes. We see that this gives the same result as when the attribute-occasion set is arranged in the vertical supermatrix, except that now the submatrices exhibit grouping of entities according to successive attributes rather than grouping of attributes according to successive entities. It is therefore clear that nothing new has been added by reordering of the data into entity by occasion slabs. The factor analysis of such an ordering of the data would be the same as for the second case involving the attribute-occasion set.

14.1.4 Additional Categories. In the previous section we have considered what may be regarded as the most obvious categories in sets of data

to be obtained in real life situations. These are certainly important, and it is probable that a great deal more attention will be given to the three category type of data matrix and to efficient methods of reproducing such a set of observed data with a smaller number of parameters. This would constitute a generalization of the lower rank approximation to data matrices of the two category type. However, it is already becoming clear that even the three category type of data matrix will not cover all meaningful categories encountered in important psychological research.

Let us consider a specific example. Suppose we have a questionnaire with a set of 64 items to which 18 individuals will respond. Let us assume that these entities or individuals are requested to give responses under a number of different conditions or instructions. For example, they may be asked to respond to the items as they apply to themselves, to the average person, to the ideal person, to the respondent as he would like to be, etc. One may have as many different conditions as he can invent.

Let us assume that there are eight of these conditions. Suppose the respondents are a group of psychiatric residents in a mental hospital. It may be expected that these residents are undergoing training and experience which will modify their responses to the items over a period of time for the varying conditions. Suppose, then, that these individuals are requested to repeat the 8 sets of responses to each of the 64 items on 4 different occasions at six month intervals.

Let us now review the essential characteristics of this data model. First we have sets of matrices involving 18 entities and 64 attributes or variables. For each condition on each occasion we have such a matrix. For example, if we have 8 conditions and 4 occasions, this means that we have 32 matrices of order 18 x 64.

We may also conceive of an additional categorical set which would consist of a set of instruments or evaluators. This would involve a number of different ways of evaluating or measuring each attribute for each entity under each condition and on each occasion. It is interesting to note that the instrument and the occasion categories are the basic concepts involved in traditional theories of reliability of measures. The instruments correspond to comparable form or comparable measure reliability, and the occasions correspond to consistency or stability over time.

In most of the measurement models, comparable form reliability usually involves only two instruments. These instruments may be persons, test booklets, hardware, or what not. For example, we may have a number of different raters evaluating the same individual on the same attribute for a given occasion. In the case of the occasion category, we have a special case of retest reliability which ordinarily involves only two occasions. The problem of evaluating change, for example, becomes sufficiently complicated from the model point of view even if we have only the four categories of entities, attributes, conditions, and occasions. It becomes even more complex if we include also the additional category of instruments. In any case, it is reasonable to assume that a general data model which is completely satisfactory should be prepared to handle at least a five category matrix.

Even though we cannot present a complete analysis of the more general problem, it may be worthwhile to examine the possibilities of arrangements for multicategory sets of data in two dimensional arrays which would be amenable to the conventional methods of factor analysis.

First we may summarize the possibilities with three category sets. We

indicate these by A, B, and C, respectively.

(A), (B,C)

(B), (A,C)

(C), (A,B)

Fig. 14.1.1

We see in Fig. 14.1.1 how we may arrange this set of data into three different kinds of two dimensional sets. The first set would have the A category for rows and the B and C categories for columns. The next set would have the B category for rows and the A and C categories for columns. The third set would have the C category for rows and the A and B categories for columns. A review of the previous subsection will show that this constitutes the three independent ways in which the data can be ordered in terms of two dimensional arrays. Any other arrangement would constitute repetitions of these, except for transposition or permutation of the matrices. Obviously, such operations on a matrix would not affect its basic structure, except for transposition of the basic orthonormals or permutation of rows and columns.

Suppose now we have four sets of categories such as entities, attributes, occasions, and conditions, which we designate A, B, C, and D, respectively. Fig. 14.1.2 indicates the ways in which these four categories can be arranged in two dimensional array matrices.

(A), (B,C,D)

(B), (A,C,D)

(C), (A,B,D)

(D), (A,B,C)

(A,B), (C,D)

(A,C), (B,D)

(A,D), (B,C)

Fig. 14.1.2

It will be noted that Fig. 14.1.2 is divided into two parts. The first part indicates those arrangements involving only one categorical set as rows and three categorical sets as columns. The second part of the figure shows how the two dimensional submatrices may be arranged into sets so that the rows will consist of two of the categories and the columns of the other two categories.

Note that in the first part of the figure the first arrangement has the members of the A category for rows and the B,C, and D categories for columns. The particular arrangement or permutations of the B, C, and D categories is irrelevant. It must be remembered that the notation used here does not mean that first, all of the B category columns are given, then all of the C categories, and finally the D categories. There would presumably be some hierarchical order of groupings and subgroupings, but these are irrelevant since these various hierarchical arrangements can be produced by permutations of columns.

The second arrangement has the members of the B categories for rows, and members of the A, C, and D categories, respectively, for columns. Here

again the order of A, C, and D are irrelevant as far as a factor analytic solution is concerned. The third arrangement has the C category for rows and the A, B, and D categories for columns. Finally, the fourth arrangement has the D category for rows and the A, B, and C categories for columns.

Next we consider the second part of Fig. 14.1.2 in which we have two categories for rows and two categories for columns. Here we see in the first row that the A and B categories are used for rows and the C and D categories for columns. Again, the ordering of the members of the A and B categories is irrelevant, since they may be permuted at will by a left hand permutation matrix. The second arrangement has the A and C categories for rows and the B and D categories for columns. The last arrangement in the second part of Fig. 14.1.2 has the A and D categories for rows and the B and C categories for columns.

It can be seen that no other arrangements exist involving two categories for rows and two for columns, which are not either transpositions or permutations of the matrices indicated in the lower part of Fig. 14.1.2. Any other combination of two categories, not involving A, which might be used for rows would actually constitute a transposition of those already indicated. For example, if B and C were used for rows, then A and D would be used for columns, and one would have the transpose of the third case in the second part of Fig. 14.1.2.

We see now that there are actually seven different ways, not involving transpositions or permutations, in which the four category data model can be arranged into 2 x 2 arrays. Any of these may have a conventional factor analysis performed upon it so as to get a basic structure solution or some other approximation or transformation of a basic structure solution. Any

of these arrangements admit of a lower rank approximation solution.

One relevant and interesting question is which of these arrangements is best from the point of view of providing the most parsimonious approximation to the data. It should be noted, for example, that if we have only three occasions, and we let this be indicated by D, the last arrangement in the first part of Fig. 14.1.2 has only three rows. A lower rank approximation to this matrix could obviously not be greater than two.

One criterion which could be useful in deciding what arrangement would give the greatest possibility for parsimonious description of the data would be to consider which combination of all possible seven indicated in Fig. 14.1.2 would result in a two array matrix such that its smaller dimension would be a maximum. Let us take, for example, the illustration used in the previous section in which we had 18 persons, 64 attributes, 8 conditions, and 3 occasions. Suppose we designate these as categories A, B, C, and D, respectively. Obviously, the arrangement here whose smallest dimension is a maximum would be A and C as rows and B and D as columns. Then the smaller dimension would be 8 x 18 or 144. This, however, is only one consideration in deciding what arrangement to use, and the problem of interpreting the results still remains.

Let us now see what happens if we have five categorical sets, say, A, B, C, D, and E. These may in particular be the sets of entities, attributes, conditions, evaluations, and occasions discussed in the previous example. Since there are five categorical sets, these can obviously be arranged so that one of the dimensions includes one set, and the other dimension includes the remaining four, as indicated in Fig. 14.1.3.

(A), (B,C,D,E)

(B), (A,C,D,E)

(C), (A,B,D,E)

(D), (A,B,C,E)

(E), (A,B,C,D)

Fig. 14.1.3

Here again, we can without loss of generality take the row dimension as the one having a single category. The data can also be arranged so that one dimension has two sets and the other has three sets, as in Fig. 14.1.4. It is irrelevant which dimension has the three categories and which has the two, so that without loss of generality we may take the row dimension as having the two categories.

We may now examine in greater detail the case of the one category row arrangement in Fig. 14.1.3. We take each of the categories in turn as the row dimension and the other four as the column dimensions of the two array matrix. Again, it must be remembered that the sequence of the symbols in parentheses is not relevant, that only the symbols themselves are important. One may have any hierarchical order desired. To adopt a convention, one may assume that the hierarchical order progresses from left to right so that the E category is the final or top hierarchy for the first four arrangements in Fig. 14.1.3. This implies a number of submatrices, each of which may be a supermatrix, such that each successive submatrix corresponds to each successive member of the E category.

Let us now examine the two category by three category, two dimensional array matrix arrangements indicated by Fig. 14.1.4.

(A,B),	(C,D,E)
(A,C),	(B,D,E)
(A,D),	(B,C,E)
(A,E),	(B,C,D)
(B,C),	(A,D,E)
(B,D),	(A,C,E)
(B,E),	(A,C,D)
(C,D),	(A,B,E)
(C,E),	(A,B,D)
(D,E),	(A,B,C)

Fig. 14.1.4

We can very simply set forth the rules for specifying the various arrangements by noting that the dimension having two categories can be made up by considering all possible different pairs of the two categories. This will obviously be 10, as indicated in the figure, because this is the number of five things taken two at a time. The column categories will, of course, be the three categories not represented in the row.

We now have the general problem of deciding which of these arrangements would be best for analyzing the data of three or more categorical sets in terms of a two dimensional array matrix amenable to the factor analytic techniques set forth in earlier chapters. It would doubtless be of interest to investigate the properties and to attempt interpretations of the various ways

in which the data can be laid out in a two dimensional matrix. It is probable that simplifying or unifying relationships among these various methods may be established.

14.2 Considerations of Origin

In the previous chapter we have given attention to the problem of origin of measurements and have discussed in part the problem of scale. We have seen how the basic structure of a matrix varies as we perform either a right or a left centering operation. We shall now consider the relationship of the metric problem to the multicategory data model which we have just discussed.

In general, any of the data arrangements considered in Fig. 14.1.1, 14.1.2, 14.1.3, and 14.1.4 must face the problem of metric. Whether we should have raw score measures, deviation measures, or standard deviation measures, and to what extent these separate considerations apply to submatrices within the set, must be decided. Ordinarily, if we have a three category entity-attribute-occasion matrix, we may assume that the difference in variance and mean for a given attribute over a given number of individuals from one occasion to the next would be of interest in the analysis of the data. Considering, therefore, the entity-attribute matrix for each occasion slab, we would not standardize the measures by columns. Such operations would obviously lose information as to relative changes or variations over time for the entities with respect to each of these attributes.

If we take the second arrangement in Fig. 14.1.1, in which B is the attribute category and A and C are, respectively, entities and occasions, we might well take standard deviation measures (th respect to the attributes, since these could be a number of different kinds of things which were not

measured in comparable units or from comparable origins. If the centering and scaling were done over all occasions for all the entities, we could see how relative means and variances fluctuated for a single attribute over the various conditions.

It may be that the problem of metric with respect to multicategory data matrices may best be resolved by considering those arrangements which involve only attributes and evaluators as the row categories. For those arrangements one would standardize by rows. This would allow differences in origins and variances to show up with respect to changing conditions and occasions, and would suppress the differences due to the arbitrary metric generally characteristic of attributes and evaluators. Once the data have been standardized for these arrangements, they can be rearranged according to other patterns indicated in Figs. 14.1.1 through 14.1.4.

14.3 Computational Considerations

Once of the problems which frequently arises and which has led to a great deal of confusion is that of determining under what conditions it is desirable to factor entities, and under what conditions attributes.

Suppose we have a two category array, irrespective of what the categories might be. They might be entities and attributes, as in the conventional case; they might be entities and occasions; or they might be attribures and occasions, as in the case of the single individual who now represents a complete population or universe. With any of these two dimensional matrices we may get either a major or a minor product moment matrix after applying some appropriate operations to achieve a specified metric. This may consist of either right or left centering or both; or it may also include scaling the members of a category, whether they are entities, attributes, or

conditions, by means of multiplication by a diagonal matrix on either the right or left or both.

It is certainly true that the basic structure of such matrices will be very much a function of the kinds of scaling and origin relocating operations that have been performed upon it. However, it does not seem to be generally recognized that, for any given set of metricizing operations--that is, for any given set of operations by which we apply additive and scaling constants to rows or columns or both, it does not matter whether we factor one set of categories or the other. If, for example, we have a data matrix which we standardize by persons or by columns, let us say, we may get the minor product moment of this normalized data matrix and perform a basic structure analysis on it. This will give us a factor loading matrix by means of which we can solve for a principal axis or basic structure factor score matrix as indicated in Chapter 4.

On the other hand, we can get a major product moment of this matrix and perform a basic structure analysis on it. The basic structure analysis will give us precisely the factor score matrix that we obtained previously by getting first the factor loading matrix and then solving for the factor score matrix.

Furthermore, if we use this factor score matrix we can then postmultiply the transpose of the standard score matrix by the factor score matrix, and this will yield the factor loading matrix which we obtained directly from the previous method by operating on a minor product moment of the standardized score matrix. It can readily be seen from the definition of the basic structure that this must be the case, because if we premultiply a matrix by the transpose of its left orthonormal, we must by definition

have the right orthonormal premultiplied by the basic diagonal. Conversely, if we premultiply the transpose of such a data matrix by its right orthonormal and then premultiply again by the inverse of its basic diagonal, we must get the left orthonormal.

It must be pointed out, however, that these reciprocal relationships hold only if the basic structure analysis is performed without altering the diagonal elements of the product moment matrix, whether this is major or minor. The communalities issue is very much a part of this problem. It should be emphasized that the definition of communalities has not been made sufficiently mathematical so that one can specify the relationship involved in these reciprocal types of solutions if communalities rather than unity are used in the diagonals of the correlation matrix. The various methods of approximating the communalities will influence the kinds of relationships obtained. Since any math-matical formulation of the communality model is extremely complex, involving complicated nonlinear relationships, we must conclude that the studies which have been done to compare the results of so-called obverse factor analysis with the conventional methods are not meaningful.

As a matter of fact, it is only for the basic structure solution that we can express precisely this reciprocal type of relationship. From a theoretical point of view therefore, much of the discussion about the factoring of people versus factoring of tests, or the obverse types of factor analysis, is irrelevant. However, there are some practical implications involved in deciding whether one does a direct factor analysis on the major product moment or on the standardized data matrix.

If, for example, one has many more attributes than entities, what we

have conventionally called the minor product moment of this data matrix is naturally larger than what we have called the major product moment. This is clear if we regard the natural order of the data matrix as having rows for entities and attributes for columns. We define the minor product moment as the natural order premultiplied by its transpose, and the major product moment as the natural order postmultiplied by its transpose. Then the order of the minor product moment will be the number of attributes and the order of the major product moment the number of entities.

If, for example, we have given a personality inventory of, say, 250 items to a group of 100 persons, and we wish to have a factor analysis of the individual items in the inventory, the conventional procedure would be to get the intercorrelations of these items and to do a factor analysis by one of the methods outlined in previous chapters. This, of course, would be a 250 x 250 matrix. This is a large matrix for any of the methods. Its analysis would be prohibitive with desk calculators and quite expensive with electronic computers. On the other hand, if we take the product of the data matrix postmultiplied by its transpose, we have a 100th order matrix which can be factor analyzed in perhaps one fifth of the time it takes to factor analyze one which has an order of 250.

We shall see, therefore, how we may proceed computationally with a data matrix having many more variables than entities. We shall assume that the matrix is normalized by columns so that we have means of 0 and standard deviations of unity. If we took the product of this matrix premultiplied by its transpose, and then divided by the number of cases, we would have precisely the intercorrelation matrix of the items.

On the other hand, if we took the product of the standardized data matrix postmultiplied by its transpose, we would not have a correlation matrix. Nevertheless, from such a product we may derive a factor loading matrix by one of the basic structure or principal axis methods. This procedure we shall indicate in the next section.

14.4 Obverse Factor Solution with Standard Metric

14.4.1 Computational Equations

14.1.1a Definition of Notation

$\underline{X}$ is the $\underline{N}$ x $\underline{n}$ matrix of raw measures.

$\underline{M}$ is the vector of means.

$\underline{D_\sigma^2}$ is the diagonal matrix of variances.

$\underline{G}$ is the major product moment of the standard score matrix.

$\underline{P}$ is the $\underline{N}$ x $\underline{m}$ matrix of factor scores.

$\underline{Q\,\delta^{\frac{1}{2}}}$ is the $\underline{n}$ x $\underline{m}$ matrix of factor loadings.

14.4.1b The Equations

$$M = \frac{X'\,1}{N} \tag{14.4.1}$$

$$\mu = \frac{X^{(2)'}1}{N} \tag{14.4.2}$$

$$\underline{D_\sigma^2} = D_\mu - D_M^2 \tag{14.4.3}$$

$$d = D_\sigma^{-2} \tag{14.4.4}$$

$$U = d\,M \tag{14.4.5}$$

$$V = X\,U \tag{14.4.6}$$

$$\alpha = U'\,M \tag{14.4.7}$$

$$W = V - \tfrac{1}{2}\,\alpha\,1 \tag{14.4.8}$$

$$Y'_{k.} = X'_{k.}\,d \tag{14.4.9}$$

$$G_{kj} = Y'_{k.}\,X_{j.} - W_k - W_j \tag{14.4.10}$$

$$G = P\,\delta\,P' \tag{14.4.11}$$

$$a = d^{\frac{1}{2}}\,(X'\,P) \tag{14.4.12}$$

14.4.2 Computational Instructions. We begin with the raw score data matrix. Although the factor loading matrix which we shall proceed to solve for is precisely the same as the solution we would get from the correlation matrix, we never actually calculate the standardized score matrix.

First we calculate the means of the variables as indicated in Eq. (14.4.1). Here we simply get a column vector of the column sums of the raw score matrix $\underline{X}$, divided by $\underline{N}$, the number of cases.

Next we calculate a vector each element of which is the sum of the squares of the elements of the corresponding column of $\underline{X}$ divided by $\underline{N}$. This is the $\underline{\mu}$ vector, as indicated in Eq. (14.4.2).

We then calculate the elements of the diagonal matrix, as shown in Eq. (14.4.3). This is obtained by constructing a diagonal matrix of the elements calculated in the vector of Eq. (14.4.2), and subtracting from it a diagonal matrix consisting of the squares of the elements calculated in Eq. (14.4.1). It can readily be shown that this is a diagonal matrix whose diagonal elements are the variances of the variables in the raw score data matrix $\underline{X}$.

Next we indicate the elements in a diagonal matrix by Eq. (14.4.4). Here we take the inverse of the diagonal matrix on the left of Eq. (14.4.3).

We next premultiply the column vector calculated in Eq. (14.4.1) by the diagonal matrix calculated in Eq. (14.4.4) to get the vector $\underline{U}$ indicated in Eq. (14.4.5).

Now we get a vector $\underline{V}$, as indicated in Eq. (14.4.6). This is the raw score data matrix postmultiplied by the vector $\underline{U}$ calculated in Eq. (14.4.5).

We then calculate a scalar quantity $\underline{\alpha}$, as indicated in Eq. (14.4.7). This is the minor product moment of the vectors calculated in Eqs. (14.4.1) and (14.4.5).

Next we calculate the vector $\underline{W}$ as in Eq. (14.4.8). This is obtained by subtracting from each element of $\underline{V}$ calculated in Eq. (14.4.6) one half the scalar $\underline{\alpha}$ calculated in Eq. (14.4.7).

Next we calculate a vector $\underline{Y_{k.}}$ beginning with $\underline{k} = 1$, as in Eq. (14.4.9). This vector is obtained by postmultiplying the $\underline{k}$th row of the $\underline{X}$ raw data matrix by the diagonal matrix calculated in Eq. (14.4.4).

This vector is then used to solve in turn for the elements of a matrix $\underline{G}$, which is the major product moment of the raw score matrix. As we recall, in this case it is smaller than the minor product moment. Eq. (14.4.10) shows how we calculate the elements for the $\underline{k}$th row and the $\underline{j}$th column. The

$\underline{Y}'_{k.}$ vector calculated in Eq. (14.4.9) is postmultiplied by the jth row of the $\underline{X}$ matrix in column form, and from this minor product are subtracted the kth and the jth elements of the $\underline{W}$ vector of Eq. (14.4.8). It is not necessary to calculate the scalar quantities of Eq. (14.4.10) for all values of j. We need calculate them only for values of j equal to or greater than k. This gives us the elements in and above the diagonal of the $\underline{G}$ matrix indicated in Eq. (14.4.11).

Eq. (14.4.11) shows the major product moment of the standardized score matrix as a function of the basic diagonal and basic orthonormals. To this matrix we now apply one of the principal axis solutions indicated in previous chapters. This may be carried to any number of factors desired, according to how much of the variance we want to account for, or what other criteria we may have for stopping the factoring.

The factor loading matrix $\underline{a}$ is indicated in Eq. (14.4.12). Here we see on the right hand side that first we postmultiply the transpose of the raw score matrix by the $\underline{P}$ matrix calculated from Eq. (14.4.11). This, it should be recognized, is precisely the principal axis factor score matrix for the normalized score matrix. That is, Eq. (14.4.11) gives us the product moment matrix we would have obtained if we had normalized the $\underline{X}$ matrix first by columns and then postmultiplied this normalized matrix by its transpose.

The next step in the calculation of the $\underline{a}$ matrix, as indicated in Eq. (14.4.12), is to premultiply the product in parentheses by the square root of the diagonal matrix calculated in Eq. (14.4.4).

It should be observed that the major saving in computations is achieved when the number of variables is much larger than the number of attributes.

Actually, as can be seen, the steps involved in Eqs. (14.4.1) through (14.4.10) are not, in general, more laborious than the calculation of the correlation matrix in which the minor product moment of the standard score matrix is involved. There is, however, an additional multiplication indicated in Eq. (14.4.12) in which the factor score matrix is postmultiplied into the transpose of the raw score data matrix. The computations for this operation ordinarily would not be great compared with the iterative procedures involved in the solution for the basic orthonormal and the basic diagonals of a very large correlation matrix.

14.4.3 Numerical Example. We illustrate the method with the same data matrix used in Chapter 12, even though it is a vertical matrix, so that we may compare the results with those obtained in Chapter 12, Section 4.

Table 14.4.1 gives the major product moment of the standardized data matrix. The number 108.0 at the lower left of the table is the sum of the diagonal elements. This should be equal to the product of the orders of the matrix. This is $12 \times 9 = 108$ and serves as a check on the computations.

Table 14.4.2 gives the normalized factor score matrix for the first three factors. The rows of this table are proportional to the columns of Table 12.4.2. The proportionality factor is $\sqrt{N}$ or $\sqrt{12}$.

The first row of Table 14.4.3 gives the first three basic diagonals of the correlation matrix corresponding to the data matrix. These results may be seen to agree closely with those of basic structure solutions for the same correlation matrix in previous chapters. The second row gives the number of iterations for each factor. The body of the table gives the first three principal axis factor loading vectors. These also agree closely with those solved for in previous chapters.

Table 14.4.1 - Major Product Moment Matrix of Standardized Score Matrix

23.328	0.057	4.059	-7.228	-9.764	-7.643	-3.918	-4.275	-2.775	2.719	3.560	1.879
0.057	22.494	4.107	-6.720	-7.004	-6.001	-1.449	-0.011	-1.007	-0.551	-0.088	-3.827
4.059	4.107	16.482	-3.170	-4.305	-3.899	-3.951	-5.014	-1.774	-0.363	0.790	-1.962
-7.228	-6.720	-3.170	9.573	5.200	4.282	0.655	0.054	0.899	-1.044	-2.355	-0.147
-9.764	-7.004	-4.305	5.200	10.916	5.283	1.535	1.212	0.990	-1.452	-2.201	-0.410
-7.643	-6.001	-3.899	4.282	5.283	8.155	1.251	0.813	0.827	-1.091	-1.964	-0.013
-3.918	-1.449	-3.951	0.655	1.535	1.251	4.605	1.795	0.752	-0.087	-1.160	-0.027
-4.275	-0.011	-5.014	0.054	1.212	0.813	1.795	5.048	0.831	0.087	-0.827	0.286
-2.775	-1.007	-1.774	0.899	0.990	0.827	0.752	0.831	2.342	-0.208	-0.638	-0.238
2.719	-0.551	-1.363	-1.044	-1.452	-1.091	-0.087	0.087	-0.208	0.858	1.097	1.035
2.560	-0.088	0.790	-2.355	-2.201	-1.964	-1.160	-0.827	-0.638	1.097	2.281	1.506
1.879	-3.827	-1.962	-0.147	-0.410	-0.013	-0.027	0.286	-0.238	1.035	1.506	1.918
108.000											

Table 14.4.2 - Factor Score Matrix for Three Factors

0.553	0.388	0.327	-0.328	-0.403	-0.326	-0.141	-0.130	-0.088	0.051	0.114	-0.015
-0.572	0.765	0.092	-0.063	0.004	-0.022	0.052	0.107	0.038	-0.098	-0.108	-0.195
0.178	0.206	-0.823	-0.198	-0.147	-0.090	0.202	0.318	0.050	0.143	0.052	0.109

Table 14.4.3 - Basic Diagonals, Number of Iterations and Factor Loading Matrix for Three Factors

3.748	2.050	1.341
11	7	8

0.717	0.489	0.352
0.740	0.474	0.324
0.773	0.291	0.408
0.553	-0.652	0.070
0.461	-0.747	0.181
0.517	-0.696	0.186
0.643	0.078	-0.589
0.615	0.165	-0.626
0.716	-0.037	-0.367

14.5 Mathematical Proof

We give below the proof that the computational outline above does yield the conventional solution for the principal axis factor loading matrix given in previous chapters.

Given the raw score matrix $\underline{X}$, let

$$M = \frac{X' 1}{N} \tag{14.5.1}$$

$$\mu = \frac{X^{(2)\prime} 1}{N} \tag{14.5.2}$$

where $\underline{X^{(2)}}$ means a matrix of the squared elements of $\underline{X}$.

$$D_\sigma^2 = D_\mu - D_M^2 \tag{14.5.3}$$

$$d = D_\sigma^{-2} \tag{14.5.4}$$

$$z = \left[I - \frac{1\,1'}{N} \right] X\, d^{\frac{1}{2}} \tag{14.5.5}$$

From Eq. (14.5.5)

$$z\, z' = \left[I - \frac{1\,1'}{N} \right] X\, d\, X' \left[I - \frac{1\,1'}{N} \right] \tag{14.5.6}$$

From Eq. (14.5.6)

$$z\, z' = X\, d\, X' - \frac{1\,1'\, X\, d\, X'}{N} - \frac{X\, d\, X'\, 1\,1'}{N} + \frac{1\,1'\, X\, d\, X'\, 1\,1'}{N^2} \tag{14.5.7}$$

From Eqs. (14.5.1) and (14.5.7)

$$z\, z' = X\, d\, X' - 1\, M'\, d\, X' - X\, d\, M\, 1' + 1\, M'\, d\, M\, 1' \tag{14.5.8}$$

Let

$$U = d\,M \tag{14.5.9}$$

Substituting Eq. (14.5.9) in Eq. (14.5.8)

$$z\,z' = X\,d\,X' - 1\,U'\,X' - X\,U\,1' + 1\,U'\,M\,1' \tag{14.5.10}$$

Let

$$\left.\begin{aligned} V &= X\,U \\ \alpha &= U'\,M \end{aligned}\right\} \tag{14.5.11}$$

From Eq. (14.5.11) in Eq. (14.5.10)

$$z\,z' = X\,d\,X' - 1\,V' - V\,1' + \alpha\,1\,1' \tag{14.5.12}$$

Let

$$W = V - \tfrac{1}{2}\alpha\,1 \tag{14.5.13}$$

From Eq. (14.5.13) in Eq. (14.5.12)

$$z\,z' = X\,d\,X' - 1\,W' - W\,1' \tag{14.5.14}$$

Let

$$G = z\,z' \tag{14.5.15}$$

From Eqs. (14.5.14) and (14.5.15)

$$G_{ij} = X'_{i.}\,d\,X_{j.} - W_i - W_j \tag{14.5.16}$$

Given the basic structure forms

$$G = P \delta P' \tag{14.5.17}$$

$$z = P \delta^{\frac{1}{2}} Q' \tag{14.5.18}$$

From Eqs. (14.5.5) and (14.5.18)

$$Q \delta^{\frac{1}{2}} P' = d^{\frac{1}{2}} X' \left[I - \frac{1\,1'}{N} \right] \tag{14.5.19}$$

From Eq. (14.5.19)

$$Q \delta^{\frac{1}{2}} = d^{\frac{1}{2}} X' P - X' 1\,1' P \tag{14.5.20}$$

From Eq. (14.5.19)

$$P' 1 = 0 \tag{14.5.21}$$

From Eq. (14.5.21) in Eq. (14.5.20)

$$Q \delta^{\frac{1}{2}} = d^{\frac{1}{2}} (X' P) \tag{14.5.22}$$

CHAPTER 15

THE PROBLEM OF SCALING

In Chapter 13 we considered the problem of origin or zero point as it affects factor analytic solutions. We saw that we may work with raw score matrices, deviation score matrices, or a number of combinations of these two methods. We learned that we may factor right centered and left centered matrices or both--that is, matrices which have means subtracted from the columns, those which have means subtracted from the rows, and those which have means subtracted from both rows and columns.

We saw also that we may conduct factor analytic solutions based on procedures which may subtract or add constants other than means to the rows and/or columns, and that these solutions vary according to the specific patterns for adding constants to rows or columns. We indicated that there may be rational procedures for determining what constants should be added, as in cases where natural zero points are available. We showed in Chapters 4 and 13 that the unit for scaling one attribute may not be comparable to that for scaling another, and that therefore rationales for making such scales comparable may be of interest.

15.1 Kinds of Scaling

It is clear that if we have a large set of measures, such as physiological, psychological, or other types, these may vary widely in comparability. For example, height may be measured in feet and a test score may be measured in terms of items correct on a 500-item test.

The conventional procedure, as indicated, has been to reduce all these to standard deviation measures. We have in general three types of possibilities for scaling. We may scale by entities, by attributes, or by both. In

any case it is well known that a particular factor analytic solution will depend on the scaling procedure. This is because the basic structure of a data matrix is altered in a very complicated fashion if the data matrix is multiplied by a diagonal matrix.

15.1.1 Scaling by Attributes. We have already considered in some detail the reasons why the problem of scaling by attributes arises. This, of course, means that we postmultiply a data matrix by a diagonal matrix. In the case of a scaling procedure which reduces all variables to unit standard deviations, we simply postmultiply the data matrix by the inverse of a diagonal matrix whose elements are standard deviations of the variables in whatever units they are measured. This has been the traditional method of scaling for factor analytic solutions.

It should be noted, however, that such a scaling procedure is specific to the particular sample to which it is applied. If one used such a scaling procedure on a particular sample and applied the same diagonal scaling matrix to a data matrix obtained from some other sample, he would not expect that that variances for the new sample would be unity. In general these would depart from unity to a greater or lesser degree. The fact that the normalizing scaling procedure is specific to a particular sample casts doubts on its validity.

15.1.2 Scaling by Entities. The problem of scaling by entities has not received much attention, or perhaps even been regarded as a relevant problem in factor analysis procedures. Certainly the question of origin by entities is of both theoretical and practical importance in the analysis of behavioral science data. We have seen how it arises in the case of ipsative type measures in personality scales. It also arises in the case differential

prediction problems. This we have discussed in Chapter 13. Obviously, the shifting of origins by entities or rows of the natural order data matrix has its analog in scaling by entities. This implies formally, of course, a multiplication on the left of the natural order data matrix by a diagonal matrix. The question of what sort of diagonal matrix is appropriate for a given problem depends on the particular interest of the investigator.

One may assume, for example, that certain of the entities should receive less weight than others in a factor analytic solution. It may be that because of biased sample selection it might be desirable to weight certain of the entities more and others less, to overcome the effects of bias. For example, if one had selected a group of individuals so that in general the higher scoring individuals were believed to be less well represented, compared to some target population, than those in the lower group, then the former might be given higher weightings. Therefore the diagonal elements of the left scaling matrix would be larger for the higher group than for the lower.

15.1.3 Scaling by Entries and Attributes. It is now obvious that a more general view of the scaling problem for a data matrix would involve scaling by both entities and attributes. Here the formal model includes both pre- and postmultiplication of the data matrix by diagonal matrices. One may make a rather basic distinction, however, between the types of left and right scaling matrices which might be considered. In the case of right diagonal scaling matrices one might well have both positive and negative elements in the scaling diagonal. For example, if one wishes to reverse the scale for certain personality item variables in a data matrix to change a negative stated statement to a positive form, then presumably one would use a negative element.

However, in the case of the left diagonal multiplier, it is difficult to see by what rationale one might wish to give a negative weight to a particular entity. In general, any left diagonal multiplier for the data matrix would almost certainly have all positive elements.

Since currently there is very little available on the rationale or technique of scaling data matrices by entities, and since no experimental or computational work has been carried out, we shall not pursue the matter further. We shall direct our attention to problems involved in the scaling of data matrices by attributes.

15.2 Scaling by Attributes

15.2.1 The General Problem of Scale. We have already discussed a number of considerations involved in the scaling of a data matrix by attributes, or the postmultiplication of the natural order data matrix by a diagonal matrix. We have pointed out that factor analytic results may vary considerably according to what scaling procedures are used. We have indicated that the Gordian knot is usually cut by using standardized measures. Nevertheless, it would seem desirable to have factor analytic procedures which are relatively independent of the scale. We shall now consider in more detail some of the criteria which suggest themselves in establishing scaling procedures.

15.2.2 Criteria for Scaling. One of the most obvious rationales for scaling has been previously suggested--namely, that of using natural units when they are available. We have indicated that in the case of the three category matrix in which one of the slabs is an entity-occasion matrix, the occasions regarded as attributes may already be in relative natural units. For example, the measures of a set of entities on typewriting scores for successive weeks are comparable both with respect to origin and scale. The

variation in scores from one week to the next for this group of entities is not some artifact of the method of evaluation, but may be of considerable interest in itself. Unfortunately, however, such natural units are not available for much of the data to be subjected to factor analytic solutions.

We have already mentioned the possibility that the factor analytic procedure may be such that the solution is relatively independent of any scaling diagonal matrix. We shall now consider some of these solutions.

15.3 The Communality Problem and Scaling

Throughout the previous chapters dealing with specific methods of factor analysis we have referred to the communality problem without being very specific as to what is meant by the term communality. True, it is defined both theoretically and computationally in texts on factor analysis. In general it is said to be that part of the variance of a system which is common to two or more variables. This is not a very precise definition.

The communality problem has also been discussed from a computational point of view. Here the problem is to determine the diagonal elements of a correlation or covariance matrix so as to reduce the rank of the matrix. To solve this problem we must decide whether we want to reduce the rank of an experimental correlation matrix precisely, or whether we want to reduce the rank of another matrix which resembles the original correlation matrix as closely as possible according to some criterion. But in the latter case we have to define "as closely as possible."

The traditional approach has used approximations to the diagonal values which enable one to account more accurately for the offdiagonal elements with a smaller number of factors than is accounted for by using unity in the

diagonals. We have seen that for correlation and residual matrices one method is to substitute the largest absolute offdiagonal element in a column for the diagonal element.

These procedures, however, do not provide precise or rigorous definitions of communalities, nor do they indicate an underlying mathematical model for their determination. They are merely verbal and arithmetic procedures with little reference to their interpretation or significance for the data matrix from which the correlation matrix is derived. We have indicated in Chapter 4 that one should be able to account completely for the results of a factor analysis in terms of the original data matrix rather than in terms of the correlation matrix.

Perhaps some of the best work on the communality problem has been done by Guttman (1958), Harris (1962), and earlier by Lawley (1940) and Rao (1955). In general these investigators have been aware of the relationship of the communality problem to the scaling problem. Implicit in their work is the notion that the communality problem is really a scaling problem.

We shall therefore consider certain types of factor analytic solutions which have techniques for solving the scaling problem built into them. These are, in effect, methods which are independent of scale or in which the scaling diagonal cancels out in the mathematical model.

15.4 Characteristics of the Methods

All of the models to be considered have certain characteristics in common. First, they are all special cases of the rank reduction method; second, they are least square or basic structure solutions; third, each solves for a scaling diagonal matrix ; and fourth, they are what may be called doubly iterative solutions.

15.4.1 Special Case of Rank Reduction Method. Each of the methods to be considered is a special case of the rank reduction formula in that the removal of each factor results in a residual matrix which is of rank one less than the previous residual matrix. Furthermore, each approximation to a factor matrix is a rank reduction solution.

15.4.2 Least Square Basic Structure Solutions. All of the solutions we shall consider are basic structure or least square solutions, with respect to the scaled matrices. This point will not be elaborated here, as it will be clarified in the computational procedure and the mathematical proofs.

15.4.3 Solution for Scaling Diagonals. As implied by the previous discussion, all of the solutions to be considered solve for a scaling diagonal matrix. It is to be observed, however, that the procedure used to solve for this scaling diagonal matrix varies considerably from one method to another. In two of the models, a single scaling diagonal matrix is solved for. In the other model, the scaling diagonal matrix is different for each factor vector. In this latter model, a scaling diagonal matrix is found for each residual matrix. The solution, however, is again independent of any particular scale that we start with, such as in the normalized data matrix.

15.4.4 Doubly Iterative Type Solutions. All of the methods to be considered might be regarded as doubly iterative, because not only does one iterate to the solution for a factor vector or matrix, but one also iterates to the scaling diagonal. This is because the scaling diagonal matrix is itself a function of the factor loading vectors, which in turn are a function of the scaling diagonal matrix. One of the consequences is that the solution may be very laborious and costly. Even with high speed computers, the cost and time may be excessive if the number of variables or attributes is large.

15.5 Kinds of Solutions

We shall consider six different kinds of solutions which are independent of scale. These may be divided into two general classes.

The first of these classes we call the specificity type solutions. The model on which these solutions are based was first proposed by Lawley (1940) and later developed in essentially the same form, but from a somewhat different set of hypotheses and assumptions, by Rao (1955).

The second class of solutions may be called the communality type solutions. These are based on a general model developed by Horst. Beginning in 1950, the method was presented in lecture notes at the University of Washington but these were not published. More recently, Kaiser, in personal communication and in conference presentations, proposed a related type of procedures.

Both the specificity and the communality types of solutions may be divided into three different variations. The first of these we shall call the successive factor method. It requires the solution of a single factor vector at a time. With the solution for each factor vector, a residual matrix is calculated and another factor vector is calculated from the residual matrix. This type of solution is analogous to the single factor residual solution outlined for the centroid and the basic structure or principal axis methods in Chapters 5 and 7, respectively. With the solution for each factor one obtains a scaling diagonal matrix which is a function of the elements of the factor vector itself.

The second type of solution for both the specificity and the communality models may be called the factor matrix solution. Here one makes some assumption as to the number of factors in the set and begins with some crude

approximation matrix of this order for the factor matrix. By a process of successive iterations one converges to the factor loading matrix, and to the scaling diagonal which is a function of all of the factor vectors.

These two types of solutions do not in general give the same results for a specified number of factors. The scaling varies from one factor vector solution to the next in the residual method, whereas it converges to a single diagonal matrix when one iterates to all the desired factor vectors simultaneously.

There is a variation of the factor matrix method which combines the features of both of the others. This we shall call the progressive factor matrix method. Here one begins with the solution for a single factor and then successively adds factors to the factor loading matrix without ever computing residual matrices.

15.6 Specificity Successive Factor Solution

We shall first take up the specificity scaling method for each of the three variations: the successive factor, the matrix, and the progressive matrix solutions. First we shall consider the successive factor type solution.

15.6.1 Characteristics of the Solution. This method is characterized by the fact that only one factor at a time is solved for, after which a residual matrix is calculated, the next factor loading vector is calculated from the residual, and so on.

All of the specificity types of solutions are similar with respect to the scaling unit solved for. The scaling unit is such that the variance of the rescaled variables is proportional to the reciprocal square root of their residual variances. That is, we define these residual variances as the

original variance of the variables less the amount of variance accounted for by a given factor or set of factors, depending on which type of solution is used. In the successive factor solution, the first scaling constant for each variable is proportional to the reciprocal square root of the difference between the original variance of the variables and the variance accounted for by the first factor. The first factor is then removed from the covariance matrix to yield a residual matrix. This residual matrix is then scaled in the same manner as are subsequent residual matrices.

15.6.2 Computational Equations

15.6.2a Definition of Notation

$\underline{C}$ is a correlation or covariance matrix.

$\underline{D_C}$ is a diagonal matrix of the diagonals of $\underline{C}$.

$\underline{{}_0a}$ is an arbitrary vector.

$\underline{D_{{}_ia}}$ is a diagonal matrix whose elements are from the vector $\underline{{}_ia}$.

$\underline{{}_iP}$ is a tolerance limit.

15.6.2b The Equations

$${}_0a = C\,1\,(1'\,C\,1)^{-\frac{1}{2}} \tag{15.6.1}$$

$${}_0D^2 = (D_C - D^2_{{}_0a})^{-1} \tag{15.6.2}$$

$${}_0U = {}_0D^2\,{}_0a \tag{15.6.3}$$

$${}_0W = C\,{}_0U - {}_0a \tag{15.6.4}$$

$$\alpha_0 = \frac{1}{\sqrt{{}_0W' \; {}_0U}} \tag{15.6.5}$$

$$ {}_1a = {}_0W \, \alpha_0 \tag{15.6.6}$$

$$ {}_iD^2 = (D_C - D^2_{{}_ia})^{-1} \tag{15.6.7}$$

$$ {}_iU = {}_iD^2 \; {}_ia \tag{15.6.8}$$

$$ {}_iW = C \; {}_iU - {}_ia \tag{15.6.9}$$

$$\alpha_i = \frac{1}{\sqrt{{}_iW' \; {}_iU}} \tag{15.6.10}$$

$$ {}_{i+1}a = {}_iW \, \alpha_i \tag{15.6.11}$$

$$\alpha_{i+1} - \alpha_i = {}_iP \tag{15.6.12}$$

$$ {}_{i+1}a = a_{.1} \tag{15.6.13}$$

$$ {}_2C = C - a_{.1} \, a'_{.1} \tag{15.6.14}$$

15.6.3 Computational Instructions. In this procedure, as in all of the methods in this section, one may begin with either a correlation matrix or a covariance matrix scaled in any convenient fashion. Ordinarily it is probably best to work with correlation matrices. These are familiar to most investigators and are convenient from the point of view of number of digits carried in the elements of the matrix.

All of the diagonals are unity, of course, in the correlation matrix. In any case, for all of the methods to be discussed, unity is used in the diagonals of correlation matrices, and variances are used in the diagonals of covariance matrices.

We begin with some arbitrary approximation to a first factor loading vector. This can be, for example, a first centroid vector, as indicated in Eq. (15.6.1). It could also be a principal axis factor loading vector.

The next step is indicated in Eq. (15.6.2). Here one gets the difference between the diagonal elements of the covariance or correlation matrix and a diagonal made up of the squared elements of the vector in Eq. (15.6.1). This is indicated on the right hand side of the equation. This diagonal matrix is then inverted to give the diagonal matrix on the left of the equation. It will be recognized that this matrix on the left is a diagonal matrix of the reciprocal of the difference between two diagonal matrices, the first of which is a diagonal matrix of variances, and the second of which is a diagonal matrix of the variances accounted for by the first approximation factor.

The next step is indicated in Eq. (15.6.3). Here we calculate a vector ${}_{\underline{0}}U$ on the right of the equation. It is obtained by premultiplying the vector of Eq. (15.6.1) by the diagonal matrix of Eq. (15.6.2).

Next we calculate the ${}_{\underline{0}}W$ vector in Eq. (15.6.4). This, as shown on the right of the equation, is obtained by postmultiplying the covariance or correlation matrix by the ${}_{\underline{0}}U$ vector calculated in Eq. (15.6.3) and subtracting from the product the vector ${}_{\underline{0}}a$ calculated in Eq. (15.6.1).

Now we calculate the scalar quantity indicated by Eq. (15.6.5). Here we get the minor product of the vectors calculated in Eqs. (15.6.3) and

(15.6.4) and take the reciprocal square root of this product.

Then we get the first rank reduction approximation to the first factor loading vector, as indicated in Eq. (15.6.6) by the vector ${}_1\underline{a}$ on the left of the equation. This is seen to be the vector ${}_0\underline{W}$ of Eq. (15.6.4) multiplied by the scalar quantity of $\underline{\alpha_0}$ of Eq. (15.6.5).

Eq. (15.6.7) gives the $\underline{i}$th approximation to the $\underline{D}^2$ matrix as the inverse of a matrix obtained by subtracting from the diagonal of the $\underline{C}$ matrix the corresponding squared elements of the current approximation to the factor loading vector. As will be seen, therefore, Eq. (15.6.7) gives a diagonal matrix which is an approximation to the inverse of the diagonal of the residual matrix.

The general equation for the $\underline{U}$ vector is given in Eq. (15.6.8). This is simply the current approximation to the factor loading vector premultiplied by the diagonal matrix of Eq. (15.6.7).

The general equation for the $\underline{W}$ vector is given by Eq. (15.6.9). This is obtained by postmultiplying the correlation matrix by the $\underline{U}$ vector of Eq. (15.6.8) and subtracting from the product the previous approximation to the factor vector.

The $\underline{i}$th approximation for the scalar quantity $\underline{\alpha}$ is the reciprocal square root of the minor product of the vectors given by Eqs. (15.6.8) and (15.6.9), as indicated on the right hand side of Eq. (15.6.10).

The general equation for the $\underline{i}$+1 approximation to the first factor loading vector is given by Eq. (15.6.11). This is the $\underline{W}$ vector of Eq. (15.6.9) multiplied by the scalar of Eq. (15.6.10).

To determine whether we have gone far enough in our approximation, we can compare successive approximations to the $\underline{a}$ vector given by Eq. (15.6.11).

However, it is probably simpler to use the criterion indicated by Eq. (15.6.12). This is the difference between successive $\underline{\alpha}$ values. These $\underline{\alpha}$ values should, in general, increase in magnitude or stabilize so that when the $\underline{P}$ value indicated by Eq. (15.6.12) is sufficiently small, we may stop the iterations for the first factor vector.

When the iterations are sufficiently close, we may regard the $\underline{i}$+1 approximations to $\underline{a}$ as the first factor loading vector, namely, $\underline{a_{.1}}$, as given in Eq. (15.6.13).

Next we calculate a residual matrix $_{\underline{2}}C$ as indicated in Eq. (15.6.14). This is obtained by subtracting the major product moment of the factor loading vector from the covariance or correlation matrix.

We now proceed through the same set of computations outlined in Eqs. (15.6.1) through (15.6.13), except that these are performed on the residual matrix given by the left side of Eq. (15.6.14), rather than on the original matrix.

Each successive residual matrix is calculated as in Eq. (15.6.15). Then the routine outlined in Eqs. (15.6.1) through (15.6.14) is applied to each of the residual matrices. The criterion of when to stop factoring may be one of those suggested in previous chapters.

15.6.4 Numerical Example. A numerical example of the method is given below. We use the same correlation matrix as in previous chapters. This correlation matrix is repeated for convenience in Table 15.6.1. The arbitrary vector for each of the four factors was taken as the unit vector. The solution is doubtless dependent on the arbitrary vectors, and currently no "best" method is available for determining these vectors.

The first row of Table 15.6.2 gives the number of iterations for each of the first four factors. The second row gives the variance accounted for by each factor. The body of the table gives the first four factor loading vectors. As in the methods of previous chapters, only the first three factors appear "significant."

It is interesting to note that the factor loading vectors bear little resemblance to the principal axis factors of Chapters 8, 9, and 10. As a matter of fact, they resemble more closely the factors given by the group centroid methods of Chapter 6. It is not clear, however, to what extent the factors might change if vastly more iterations were taken.

15.7 The Specificity Factor Matrix Solution

15.7.1 Characteristics of the Method. In this method the residual variance scaling matrix is based on all of the factors to be solved for, rather than on a single factor as in the method just outlined. Therefore we do not have a rescaling after each factor vector. The method is different also in that, instead of solving for a single factor at a time and getting a residual matrix for each cycle, we start with a rough approximation to the complete factor matrix in which some specified number of factors is assumed. We then iterate successively to the factor loading matrix and to the scaling diagonal whose elements are the reciprocal square roots of the residual variances.

15.7.2 Computational Equations

15.7.2a Definition of Notation

$\underline{C}$ is a covariance matrix.

$\underline{D}_{\underline{C}}$ is the diagonal matrix from $\underline{C}$.

Table 15.6.1 - The Correlation Matrix

1.000	0.829	0.768	0.108	0.033	0.108	0.298	0.309	0.351
0.829	1.000	0.775	0.115	0.061	0.125	0.323	0.347	0.369
0.768	0.775	1.000	0.272	0.205	0.238	0.296	0.271	0.385
0.108	0.115	0.272	1.000	0.636	0.626	0.249	0.183	0.369
0.033	0.061	0.205	0.636	1.000	0.709	0.138	0.091	0.254
0.108	0.125	0.238	0.626	0.709	1.000	0.190	0.103	0.291
0.298	0.323	0.296	0.249	0.138	0.190	1.000	0.654	0.527
0.309	0.347	0.271	0.183	0.091	0.103	0.654	1.000	0.541
0.351	0.369	0.385	0.369	0.254	0.291	0.527	0.541	1.000

Table 15.6.2 - Specificity Successive Factor Method. Number of Iterations, Variance Accounted for, and First Four Factor Vectors

24	21	30	30
2.9730	2.0250	1.1978	0.0898
0.8910	-0.1076	-0.0480	-0.0116
0.9067	-0.0867	-0.0182	-0.0654
0.8572	0.0914	-0.0932	0.0874
0.2151	0.7438	0.0475	0.2218
0.1389	0.8266	-0.0821	-0.0500
0.2025	0.7968	-0.0680	-0.0587
0.3911	0.1770	0.6572	0.0118
0.3957	0.1012	0.7303	-0.0627
0.4539	0.2908	0.4556	0.1360

${}_{0}a$ is an arbitrary factor matrix approximation of specified width.

${}_{i}a$ is the *i*th approximation to the factor matrix.

$D_{{}_{i}a\,{}_{i}a'}$ is the diagonal matrix of ${}_{i}a\,{}_{i}a'$.

${}_{i}t$ is a triangular matrix.

P is a tolerance limit.

15.7.2b The Equations

$${}_{0}D^{2} = (D_{C} - D_{{}_{0}a\;{}_{0}a'})^{-1} \qquad (15.7.1)$$

$${}_{0}U = {}_{0}D^{2}\;{}_{0}a \qquad (15.7.2)$$

$${}_{0}W = C\;{}_{0}U - {}_{0}a \qquad (15.7.3)$$

$$S = {}_{0}W'\;{}_{0}U \qquad (15.7.4)$$

$$\begin{bmatrix} {}_{0}t \\ {}_{0}W\;{}_{0}t'^{-1} \end{bmatrix} {}_{0}t' = \begin{bmatrix} {}_{0}S \\ {}_{0}W \end{bmatrix} \qquad (15.7.5)$$

$${}_{1}a = {}_{0}W\;{}_{0}t'^{-1} \qquad (15.7.6)$$

$${}_{1}D^{2} = (D_{C} - D_{{}_{1}a\;{}_{1}a'})^{-1} \qquad (15.7.7)$$

$$ {}_iU = {}_iD^2 \, {}_ia \tag{15.7.8} $$

$$ {}_iW = C \, {}_iU - {}_ia \tag{15.7.9} $$

$$ {}_iS = {}_iW' \, {}_iU \tag{15.7.10} $$

$$ \begin{bmatrix} {}_it & \\ {}_iW & {}_it'^{-1} \end{bmatrix} {}_it' = \begin{bmatrix} {}_iS \\ {}_iW \end{bmatrix} \tag{15.7.11} $$

$$ {}_{i+1}a = {}_iW \, {}_it'^{-1} \tag{15.7.12} $$

$$ H = \left| \frac{tr\,({}_it)}{tr\,({}_{i+1}t)} \right| - P \tag{15.7.13} $$

15.7.3 Computational Instructions. In this variation of the specificity method we postulate a given number of factors and begin with an arbitrary factor loading matrix including the assumed number of factors. This we may obtain from the methods of previous chapters.

We first calculate a diagonal matrix as in Eq. (15.7.1). This is obtained by subtracting from the diagonal of the covariance matrix the diagonal of the major product moment of the arbitrary factor loading matrix. Then we take the inverse of this difference matrix as indicated on the right of Eq. (15.7.1).

The next step is to calculate the $\underline{U}$ matrix, as indicated in Eq. (15.7.2). Here we premultiply the first approximation to the factor loading matrix by the diagonal matrix calculated in Eq. (15.7.1).

We then calculate a $\underline{W}$ matrix, as in Eq. (15.7.3). This is obtained by

postmultiplying the covariance matrix $\underline{C}$ by the $\underline{U}$ matrix calculated in Eq. (15.7.2) and subtracting from the product the arbitrary approximation to the factor matrix.

Next we calculate the matrix $\underline{S}$ in Eq. (15.7.4). This is the minor product of the matrices calculated in Eqs. (15.7.2) and (15.7.3). It can be seen by the definitions of these matrices that the product is symmetric.

We now indicate a supermatrix of the matrices solved for in Eqs. (15.7.3) and (15.7.4). This is given in the right hand side of Eq. (15.7.5). The left hand side of Eq. (15.7.5) indicates a partial triangular factoring of the supermatrix.

The lower part of the left partial triangular factor is then the first approximation to the factor loading matrix, as indicated on the right of Eq. (15.7.6).

The general equations are given in Eqs. (15.7.7) through (15.7.13). Eq. (15.7.7) gives the general equation for the $\underline{D}^2$ matrix. This, as indicated on the right, is obtained by subtracting from the diagonal of the covariance matrix,the diagonal of the major product moment of the current approximation to the factor loading matrix,and then taking the inverse of this difference diagonal matrix.

Eq. (15.7.8) gives the $\underline{i}$th approximation to the $\underline{U}$ matrix, which is the current approximation to the factor loading matrix premultiplied by the diagonal matrix of Eq. (15.7.7).

The $\underline{i}$th approximation to the $\underline{W}$ matrix is given by Eq. (15.7.9). This is the product of the covariance matrix postmultiplied by the $\underline{U}$ matrix calculated in Eq. (15.7.8), less the previous approximation to the factor matrix.

Eq. (15.7.10) indicates the $\underline{i}$th approximation to the symmetric $\underline{S}$ matrix, which is the minor product of the $\underline{U}$ and $\underline{W}$ matrices calculated, respectively, in Eqs. (15.7.8) and (15.7.9).

We indicate in general the supermatrix made up of the matrices calculated in Eqs. (15.7.9) and (15.7.10), as on the right hand side of Eq. (15.7.11). We then indicate the partial triangular factoring of this supermatrix, as shown on the left side of Eq. (15.7.11).

The lower matric element of the supermatrix on the left hand side of Eq. (15.7.11) gives the next approximation to the factor loading matrix, as indicated in Eq. (15.7.12). It can be proved that the triangular matrix indicated in the upper element of the left hand matrix in Eq. (15.7.11) converges to a diagonal matrix whose elements are the largest roots or basic diagonal elements of the scaled covariance matrix.

We then assume that the traces of successive $\underline{t}$ matrices or the sums of their diagonal elements will converge to some value. Therefore, as indicated in Eq. (15.7.13), we take the ratios of successive traces to get $\underline{H}$ values. When these are sufficiently close to unity the computations cease.

15.7.4 Numerical Example. We use the same correlation matrix as in the previous methods. For the aabitrary matrix we take the first four principal axis factor vectors of this matrix as found in previous solutions.

For convenient reference the first four principal axis row vectors are given in Table 15.7.1.

Table 15.7.2 gives the successive traces of the $\underline{t}$ matrices for thirty iterations.

The first row of Table 15.7.3 gives the variance accounted for by each

of the first four factors. The body of the table gives the first four column factor loading vectors. These factor loadings bear little relation to the principal axis factors of Table 15.7.1. However, again it is apparent that the fourth factor may be ignored. While the factor loadings are not the same within decimal error as those of Table 15.6.3, we may compare factors with the three highest loadings. For Tables 15.6.3 and 15.7.3 we have as comparable factors respectively, factors 1 and 2, 2 and 3, 3 and 1. Again it may be that a great many more iterations would yield a matrix considerably different from that of Table 15.7.3.

15.8 The Specificity Progressive Factor Matrix Method

15.8.1 Characteristics of the Method. This method is essentially a combination of the previous two methods. It uses the same scaling rationale--that is, the reciprocal square roots of the residual variances of the attributes. It starts with a single factor and proceeds by adding successive factors.

It differs essentially from the first method, however, in that no residual matrices are calculated. It is similar to the second method in that only a single scaling of the variables is solved for. It differs in that no assumptions are made as to the number of factors required.

15.8.2 The Computational Equations

15.8.2a Definition of Notation

$\underline{(k)}$ subscript designates a matrix of width $\underline{k}$.

Other notation is the same as in Section 15.7.2a.

Table 15.7.1 - First Four Principal Axis Row Factor Vectors of the Correlation Matrix

0.717	0.740	0.773	0.556	0.463	0.518	0.640	0.615	0.715
0.493	0.478	0.296	-0.649	-0.744	-0.694	0.080	0.166	-0.034
0.350	0.322	0.406	0.068	0.181	0.188	-0.588	-0.621	-0.369
0.030	-0.056	0.009	0.102	-0.115	-0.129	-0.288	-0.165	0.566

Table 15.7.2 - Traces of Successive $\underline{t}$ Matrices for Thirty Iterations

67.3460	54.1334	54.0273
57.4669	54.1246	54.0201
55.3702	0.1747	54.0130
54.6664	0.1646	54.0061
54.3994	0.2453	
54.2859	0.4108	
54.2323	0.2689	
54.2035	0.3148	
54.1855	0.3076	
54.1725	0.3691	
54.1616	0.0314	
54.1518	54.0421	
54.1425	54.0346	

Table 15.7.3 - Specificity Factor Matrix Method. Variance Accounted for and Column Factor Vectors for First Four Factors

2.8295	1.6250	1.6337	0.6243
0.5091	0.7508	-0.0229	0.0439
0.5297	0.7447	-0.0087	0.0093
0.5390	0.6577	0.1617	0.0735
0.4148	-0.1135	0.6337	-0.0521
0.3020	-0.1292	0.7893	-0.0128
0.3476	-0.0770	0.7470	-0.0199
0.5858	0.0322	-0.0403	-0.5875
0.5901	0.0384	-0.1302	-0.5152
0.9611	-0.1897	-0.0760	0.0559

15.8.2b **The Equations**

$${}_0a_{(k-1)} = (a_{.1} \cdots a_{.k-1}) \qquad (15.8.1)$$

$${}_0a_{(k)} = (a_{(k-1)}\, ,\, {}_0a) \qquad (15.8.2)$$

$${}_0D_k^2 = (D_C - D_{{}_0a(k)\,{}_0a'(k)})^{-1} \qquad (15.8.3)$$

$${}_0U_{(k)} = {}_0D_k^2\,{}_0a_{(k)} \qquad (15.8.4)$$

$${}_0W_{(k)} = C\,{}_0U_{(k)} - {}_0a_{(k)} \qquad (15.8.5)$$

$${}_0t_{(k)}\,{}_0t'_{(k)} = {}_0W'_{(k)}\,{}_0U_{(k)} \qquad (15.8.6)$$

$${}_1a_{(k)} = {}_0W_{(k)}\,{}_0t'^{-1}_{(k)} \qquad (15.8.7)$$

$${}_iD_k^2 = (D_C - D_{{}_ia(k)\,{}_ia'(k)})^{-1} \qquad (15.8.8)$$

$${}_iU_{(k)} = {}_iD_k^2\,{}_ia_{(k)} \qquad (15.8.9)$$

$${}_iW_{(k)} = C\,{}_iU_{(k)} - {}_ia_{(k)} \qquad (15.8.10)$$

$${}_it_{(k)}\,{}_it'_{(k)} = {}_iW'_{(k)}\,{}_iU_{(k)} \qquad (15.8.11)$$

$${}_{i+1}a_{(k)} = {}_iW_{(k)}\,{}_it'^{-1}_{(k)} \qquad (15.8.12)$$

15.8.3 Computational Instructions. This method begins with an arbitrary vector as in the first specificity type of solution. The method for getting the first factor vector is the same as in that solution.

We next proceed to indicate a factor loading matrix, as indicated in Eq. (15.8.1), where now $\underline{k}$-1 is the number of factors currently solved for.

We then indicate an augmented matrix to which one more factor has been added, as indicated in Eq. (15.8.2). To begin with, the first matrix in the parentheses on the right of Eq. (15.8.2) is simply the first factor loading vector $\underline{a}_{.1}$. This is augmented now by a second arbitrary vector which may be assumed to be a reasonable approximation to the second factor loading vector.

We then have, as in Eq. (15.8.3), a diagonal matrix which is, as indicated on the right, the reciprocal of the diagonal of the covariance matrix minus the diagonal of the major product moment of the matrix in Eq. (15.8.2).

We indicate in Eq. (15.8.4) the matrix of Eq. (15.8.2) premultiplied by the diagonal matrix of Eq. (15.8.3).

Eq. (15.8.5) is obtained by premultiplying the matrix of Eq. (15.8.4) by the covariance matrix and subtracting the arbitrary approximation to the factor matrix from it.

Eq. (15.8.6) indicates the minor product moment of the matrices of Eqs. (15.8.4) and (15.8.5) as a major product of a partial triangular matrix. In particular, this could be solved for by means of the partial triangular factoring of the supermatrix indicated in Eq. (15.7.5) of the previous method.

Eq. (15.8.7) gives the first approximation to the factor loading matrix of width $\underline{k}$ as the $\underline{W}$ matrix of Eq. (15.8.5) postmultiplied by the inverse of the upper triangular matrix of Eq. (15.8.6).

The general iterative type of solution is indicated by Eqs. (15.8.8) through (15.8.12). Here the equations are, respectively, the same as Eqs. (15.8.3) through (15.8.7), except that now the prescript becomes $\underline{i}$ for the

$\underline{i}$th approximation. In this type of solution we may again iterate to some convergence criterion for the trace of the triangular matrix, indicated on the left of Eq. (15.8.11). If the traces of two successitve $\underline{t}$ matrices are sufficiently close, we may assume that the approximation is sufficiently close for the current number of factors, $\underline{k}$.

Once this criterion has been satisfied, we again augment the currently stabilized factor loading matrix by another arbitrary vector, which is presumably reasonably orthogonal to the current factor vectors and which is not too poor an approximation to the next factor vector we wish to obtain.

We then proceed again through Eqs. (15.8.2) to (15.8.7) to get a first approximation to the factor loading matrix with one more factor added.

Going through Eqs. (15.8.8) through (15.8.12), we continue to iterate, increasing the value of subscript $\underline{i}$ until the solution has stabilized to some specified tolerance with reference to the traces of two successive $\underline{t}$ matrices.

We proceed to augment the matrix in Eq. (15.8.2) until we have accounted for enough factors, according to some specified criterion. This criterion may well be simply the sums of squares of elements of a currently stabilized $\underline{a}$ matrix, such as given in Eq. (15.8.12). The sums of squares of these elements are, of course, the amount of variance accounted for by the given number of factors.

15.8.4 Numerical Example. We use the same correlation matrix as in the previous section. In this numerical example we use as the arbitrary vector for each new factor the corresponding principal axis vector of the correlation matrix.

The first row of Table 15.8.1 gives the variance accounted for by each of the first four factors. The body of the table gives the first four factor vectors. Again it appears that the fourth factor may be ignored.

Table 15.8.1 - Specificity Progressive Factor Matrix Method. Variance Accounted for and First Four Factor Vectors

3.3946	1.8123	1.0936	0.2634
0.8251	-0.3481	-0.1503	-0.0012
0.8444	-0.3236	-0.1332	0.0332
0.8203	-0.1577	-0.2405	-0.0311
0.3659	0.6436	-0.2086	0.0083
0.2814	0.7159	-0.3664	0.0805
0.3410	0.6670	-0.3468	0.0579
0.5341	0.2126	0.4568	0.2343
0.5445	0.1573	0.5665	0.2987
0.6458	0.3463	0.4095	-0.3276

15.9 The Communality Successive Factor Method

We shall now consider the first of the communality class of scaling methods. In these methods we have the three different types of solutions--namely, the successive factor vector solution, the factor matrix solution, and the progressive factor matrix solution. These methods are essentially the same as the specificity scaling methods except that the scaling diagonal is different, and no diagonal matrix is subtracted from the correlation matrix.

Here the scaling constants are inversely proportional to the square roots of the variances accounted for by the vectors solved for. This principal of scaling is just the opposite of that used in the specificity method. In the specificity method the scaling is such that the variance unaccounted for by the factors is the same for all variables, while in the communality method the scaling is such that the variance accounted for by the factors is the same for all variables. In this latter procedure it is assumed that more weight should be given to the variables which otherwise would have less of their variance accounted for by the factors.

We begin now with the computational equations for the successive factor method.

15.9.1 The Computational Equations

15.9.1a Definition of Notation

$\underline{C}$ is a correlation or covariance matrix.

$\underline{V}$ is an arbitrary vector.

$\underline{{}_ia}$ is the $\underline{i}$th approximation to a factor vector.

$\underline{D_{{}_ia}}$ is a diagonal matrix of the elements of $\underline{{}_ia}$.

$\underline{H_i}$ is a tolerance limit.

$\underline{{}_{k+1}C}$ is the $\underline{k}$th residual matrix.

15.9.1b The Equations

$$ {}_0a = C\,V\,(V'\,C\,V)^{-\frac{1}{2}} \tag{15.9.1}$$

$$ {}_0D = D^{-1}\,{}_0a \tag{15.9.2}$$

$$ {}_0U = {}_0D\,1 \tag{15.9.3}$$

$$ {}_0W = C\,{}_0U \tag{15.9.4}$$

$$ \alpha_0 = \frac{1}{\sqrt{{}_0W'\,{}_0U}} \tag{15.9.5}$$

$$ {}_1a = {}_0W\,\alpha_0 \tag{15.9.6}$$

$$ {}_iD = D^{-1}\,{}_ia \tag{15.9.7}$$

$$ {}_iU = {}_iD\,1 \tag{15.9.8}$$

$$ {}_iW = C\,{}_iU \tag{15.9.9}$$

$$ \alpha_i = \frac{1}{\sqrt{{}_iW'\,{}_iU}} \tag{15.9.10}$$

$$ {}_{i+1}a = {}_iW\,\alpha_i \tag{15.9.11}$$

$$\left| \frac{\alpha_i}{\alpha_{i+1}} - 1 \right| = H_i \tag{15.9.12}$$

$$a_{.i} = {}_{i+1}a \tag{15.9.13}$$

$$_2C = C - a_{.1}\, a'_{.1} \tag{15.9.14}$$

$$_{k+1}C = {}_kC - a_{.k}\, a'_{.k} \tag{15.9.15}$$

15.9.2 Computational Instructions. The computational procedure for this method is the same as for the corresponding method in the specificity class of solutions, except that the $\underline{D}^2$ matrix and the $\underline{W}$ vector are calculated differently.

By means of an arbitrary vector $\underline{V}$ we first calculate the rank reduction $\underline{_0a}$ vector as in Eq. (15.9.1).

Next we take the inverse of the elements of the $\underline{_0a}$ matrix given in Eq. (15.9.1) to construct the $\underline{D}$ matrix given in Eq. (15.9.2). This is a diagonal matrix of the inverse of the elements in the vector given by Eq. (15.9.1).

Eq. (15.9.3) indicates $\underline{U}$ as a vector of the elements of the $\underline{D}$ matrix given by Eq. (15.9.2).

Eq. (15.9.4) is the $\underline{C}$ matrix postmultiplied by the $\underline{U}$ vector of Eq. (15.9.3).

Eq. (15.9.5) is a scalar quantity which is the reciprocal square root of the minor product of the vectors of Eqs. (15.9.3) and (15.9.4).

Eq. (15.9.6) gives the $\underline{W}$ vector of Eq. (15.9.4) multiplied by the scalar of Eq. (15.9.5).

Eqs. (15.9.7) through (15.9.11) indicate the iterations as in the analogous specificity method.

Eq. (15.9.12) indicates the tolerance limit which is assumed to give a sufficiently close approximation.

The $\underline{i}$+1 approximation to $\underline{a}_{.1}$ is then taken as the first factor vector, as indicated in Eq. (15.9.13).

Eq. (15.9.14) gives the first residual matrix as in the specificity method.

Eq. (15.9.15) gives a generalization of Eq. (15.9.14).

The procedures for Eqs. (15.9.1) through (15.9.13) are applied to the successive residual matrices.

15.9.3 Numerical Example. We use the same correlation matrix as in the three previous examples to illustrate this method. The unit vector is taken for the arbitrary vectors. The computations are a little simpler with respect to the $\underline{D}$ matrix, since it involves only the factor loading vector and does not involve elements from the covariance matrix.

The first row of Table 15.9.1 gives the number of iterations for each of three factors. The second row gives the variance accounted for by each factor. The body of the table gives the first three factor loading vectors.

15.10 The Communality Factor Matrix Solution

15.10.1 The Computational Equations

15.10.1a Definitions of Notation

$\underline{C}$ is a covariance or residual matrix.

$\underline{_i a}$ is the $\underline{i}$th approximation to the factor matrix.

Table 15.9.1 - Communality Successive Factor Solution. Number of Iterations, Variance Accounted for, and First Three Factors

5	30	30
3.7106	1.8969	1.3096
0.6398	0.3749	0.4982
0.6639	0.3751	0.4818
0.7117	0.2347	0.5164
0.6326	-0.5577	-0.2927
0.5685	-0.6786	0.0257
0.6087	-0.6347	0.0133
0.6351	0.3527	-0.4426
0.6053	0.4040	-0.4655
0.7002	0.3139	-0.2522

$D_{_1a\ _1a'}$ is a diagonal matrix of the diagonals of $_1a\ _1a'$.

$_1t$ is a triangular matrix.

H_1 is a tolerance limit.

15.10.1b The Equations

$$_0D^2 = D^{-1}_{_0a\ _0a'} \tag{15.10.1}$$

$$_0U = {_0D^2}\ _0a \tag{15.10.2}$$

$$_0W = C\ _0U \tag{15.10.3}$$

$$_0t\ _0t' = {_0W'}\ _0U \tag{15.10.4}$$

$$_1a = {_0W}\ _0t'^{-1} \tag{15.10.5}$$

$$_1D^2 = D^{-1}_{_1a\ _1a'} \tag{15.10.6}$$

$$_1U = {_1D^2}\ _1a \tag{15.10.7}$$

$$_1W = R\ _1U \tag{15.10.8}$$

$$_1t\ _1t' = {_1W'}\ _1U \tag{15.10.9}$$

$$_{1+1}a = {_1W}\ _1t'^{-1} \tag{15.10.10}$$

$$\left| \frac{tr\ (_1t)}{tr\ (_{1+1}t)} - 1 \right| = H_1 \tag{15.10.11}$$

15.10.2 Computational Instructions. The computational instructions for this method are almost identical to those of the corresponding specificity factor scaling method, except that again the D^2 matrices of Eqs. (15.10.1)

and (15.10.6) and the $\underline{W}$ matrices of Eqs. (15.10.3) and (15.10.8) are calculated differently. It will be seen that the $\underline{D}^2$ matrices are obtained by taking the reciprocal of the diagonal of the major product moment of the factor loading matrix, rather than by subtracting this diagonal from the original diagonal of variances. This difference in the calculation of the $\underline{D}$ matrices reflects the difference in the underlying rationale of the method. The $\underline{W}$ matrices are different in that they do not involve the subtraction of the current $\underline{a}$ matrix.

It will be noted that Eqs. (15.10.4) and (15.10.9) indicate the minor product of the $\underline{W}$ and the $\underline{U}$ matrices as the major product moment of partial triangular factors. These equations do not explicitly indicate the partial triangular factoring of a type 3 supervector, as indicated in Eqs. (15.7.5) and (15.7.11). However, the computations may be carried out in the same fashion.

15.10.3 Numerical Example. Again we take the same correlation matrix as in the previous illustrations. We also take its first four principal axis vectors as the arbitrary matrix.

Table 15.10.1 gives the first four factor loading vectors for the correlation matrix as determined from the rescaled matrix. These are considerably different from those in Table 15.9.1.

15.11 The Communality Progressive Factor Matrix Method

15.11.1 Computational Equations

15.11.1a Definition of Notation

$\underline{C}$ is the covariance or correlation matrix.

$\underline{{}_i a_{(k)}}$ is the $\underline{i}$th approximation to a factor loading matrix of width $\underline{k}$.

Table 5.10.1 - Communality Factor Matrix Method, First Four Factor Vectors

0.6927	0.5144	0.3644	-0.0312
0.7165	0.5022	0.3346	-0.0585
0.7593	0.3215	0.4158	0.0135
0.5903	-0.6276	0.0442	0.1678
0.4992	-0.7235	0.1528	-0.1334
0.5515	-0.6716	0.1633	-0.1568
0.6369	0.1303	-0.5901	-0.2628
0.6078	0.2166	-0.6236	-0.1325
0.7053	0.0158	-0.3522	0.5589

$D_{{}_i a_{(k)}\, {}_i a'_{(k)}}$ is a matrix of the diagonal of ${}_i a_{(k)}\, {}_i a'_{(k)}$.

15.11.1b The Equations

$$a_{(k-1)} = (a_{.1} - a_{.k-1}) \qquad (15.11.1)$$

$$_0a_{(k)} = (a_{(k-1)}, {}_0a) \qquad (15.11.2)$$

$$_0D^2_{(k)} = D^{-1}_{{}_0a_{(k)}\, {}_0a'_{(k)}} \qquad (15.11.3)$$

$$_0U_{(k)} = {}_0D^2_{(k)}\, {}_0a_{(k)} \qquad (15.11.4)$$

$$_0W_{(k)} = C\, {}_0U_{(k)} \qquad (15.11.5)$$

$$_0t_{(k)}\, {}_0t'_{(k)} = {}_0W'_{(k)}\, {}_0U_{(k)} \qquad (15.11.6)$$

$$_1a_{(k)} = {}_0W_{(k)}\, {}_0t'^{-1}_{(k)} \qquad (15.11.7)$$

$$_iD^2_{(k)} = D^{-1}_{{}_ia_{(k)}\, {}_ia'_{(k)}} \qquad (15.11.8)$$

$$_iU_{(k)} = {}_iD^2_{(k)}\, {}_ia_{(k)} \qquad (15.11.9)$$

$$_iW_{(k)} = C\, {}_iU_{(k)} \qquad (15.11.10)$$

$$_it_{(k)}\, {}_it'_{(k)} = {}_iW'_{(k)}\, {}_iU_{(k)} \qquad (15.11.11)$$

$$_{i+1}a_{(k)} = {}_iW_{(k)}\, {}_it'^{-1}_{(k)} \qquad (15.11.12)$$

$$\left| 1 - \frac{tr\ (t_k)}{tr\ (0)} \right| = P_k \qquad (15.11.13)$$

15.11.2 Computational Instructions. Here the computational procedure is essentially the same as that of the corresponding specificity method, except that again the $\underline{D}$ matrices are calculated only as the major product moment of the factor loading matrix, and the $\underline{W}$ matrices do not involve the subtraction of the current approximation to the factor loading matrix.

15.11.3 Numerical Example. The correlation matrix is the same as in the previous examples.

The first row of Table 15.11.1 gives the variance accounted for by each of the first three factors. The body of the table gives the three factor vectors.

Although the three communality scaling methods give different results, the general orders of magnitude of the factor loadings compare favorably with one another and with the corresponding principal axis factor loadings. The signs for corresponding elements of all four sets are the same for the first three factors.

15.12 Mathematical Proofs

15.12.1 Proof of the Specificity Successive Factor Method

Let $\underline{C}$ be the correlation of the covariance matrix, and consider

$$C - D^{-2} - a\ a' = {}_1C \qquad (15.12.1)$$

where $\underline{a}$ is a vector or matrix of specified width and

$$D = (D_C - D'_{a\ a'})^{-\frac{1}{2}} \qquad (15.12.2)$$

We let

Table 15.11.1 - Communality Progressive Factor Matrix Method. Variance Accounted for and First Three Factor Vectors

3.7343	2.0508	1.3421
0.6655	0.5042	0.4242
0.6904	0.4932	0.3969
0.7306	0.3105	0.4711
0.5884	-0.6275	0.0621
0.4908	-0.7257	0.1613
0.5429	-0.6746	0.1763
0.6630	0.1528	-0.5376
0.6331	0.2390	-0.5627
0.7476	0.0312	-0.3406

$$D\,a = \alpha \tag{15.12.3}$$

$$D\,(C - D^{-2})\,D\,\alpha = \alpha\,\alpha'\alpha \tag{15.12.4}$$

We may begin with some approximation to the first principal axis vector such as the centroid. We let this vector be ${}_0a$, and calculate

$${}_0D^2 = (D_C - D^2_{{}_0a})^{-1} \tag{15.12.5}$$

Then consider

$${}_0D\,{}_1a = {}_0D\,[C - {}_0D^{-2}]\,{}_0D^2\,{}_0a\,[{}_0a'\,{}_0D\,(C - {}_0D^{-2})\,{}_0D\,{}_0a]^{-\frac{1}{2}} \tag{15.12.6}$$

which is a rank reduction form.

We let

$${}_0U = {}_0D^2\,{}_0a \tag{15.12.7}$$

From Eqs. (15.12.6) and (15.12.7)

$${}_1a = (C - {}_0D^{-2})\,{}_0U\,[{}_0U'\,(C - {}_0D^{-2})\,{}_0U]^{-\frac{1}{2}} \tag{15.12.8}$$

If we let

$${}_0W = C\,{}_0U - {}_0a \tag{15.12.9}$$

and

$${}_0\alpha = \frac{1}{\sqrt{{}_0W'\,{}_0U}} \tag{15.12.10}$$

Then from Eqs. (15.12.8), (15.12.9), and (15.12.10),

$${}_1a = {}_0W\,{}_0\alpha \tag{15.12.11}$$

In general then, we have

$$ {}_iD^2 = (D_C - D^2_{{}_ia})^{-1} \tag{15.12.12} $$

$$ {}_iU = {}_iD^2\, {}_ia \tag{15.12.13} $$

$$ {}_iW = C\, {}_iU - {}_ia \tag{15.12.14} $$

$$ {}_i\alpha = \frac{1}{\sqrt{{}_iW'\, {}_iU}} \tag{15.12.15} $$

$$ {}_{i+1}a = {}_iW\, {}_i\alpha \tag{15.12.16} $$

We may continue until ${}_i\alpha$ stabilizes and then calculate the residual

$$ {}_1C = C - a_{.1}\, a'_{.1} \tag{15.12.17} $$

The operations on ${}_1C$ are the same as for C. Successive residual ${}_iC$'s may be obtained to a specified number of factors.

We show in Section 15.12.7 that the solution is independent of scale.

15.12.2 Proof of the Specificity Factor Matrix Method

Let a be a factor loading matrix of specified width. We may still use Eqs. (15.12.1) through (15.12.7) without loss of generality. We now, however, introduce

$$ {}_0t\, {}_0t' = {}_0W'\, {}_0U \tag{15.12.18} $$

Analogous to Eq. (15.12.11), we now write

$$ {}_1a = {}_0W\, {}_0t'^{-1} \tag{15.12.19} $$

In general then, for $\underline{a}$ of any width, we have

$$ {}_iD^2 = (D_C - D\,{}_ia\,{}_ia')^{-1} \tag{15.12.20} $$

$$ {}_iU = {}_iD^2\,{}_ia \tag{15.12.21} $$

$$ {}_iW = C\,{}_iU - {}_ia \tag{15.12.22} $$

$$ {}_it\,{}_it' = {}_iW'\,{}_iU \tag{15.12.23} $$

$$ {}_{i+1}a = {}_iW\,{}_it'^{-1} \tag{15.12.24} $$

These iterations may continue until

$$ \mathrm{tr}\,({}_{i+1}t) - \mathrm{tr}\,({}_it) = P \tag{15.12.25} $$

for $\underline{P}$ sufficiently small. Then $\underline{{}_it}$ will approach the basic diagonal $\underline{\delta}$ of

$$ D\,C\,D - I = Q\,\delta\,Q' \tag{15.12.26} $$

and

$$ Q\,\delta^{\frac{1}{2}} = D\,a = \alpha \tag{15.12.27} $$

That this solution is independent of scale is shown in Section 15.12.7.

15.12.3 Proof of the Specificity Progressive Factor Matrix Method

Let Eqs. (15.12.20) through (15.12.24) be the iterative procedure for $\underline{a}$ of width $\underline{k}$. In particular, $\underline{k}$ may be 1. Continue until Eq. (15.12.25) is satisfied and indicate

$$ a_{(k)} = (a_{.1} \cdots a_{.k}) \tag{15.12.28} $$

Let

$$ {}_0a(k+1) = (a_{,1} \ldots a_{,k}, {}_0a) \qquad (15.12.29)$$

where $\underline{{}_0a}$ is an arbitrary vector distinct from all the preceding $\underline{a_{,i}}$'s and preferably orthogonal to them.

Then let

$$ {}_0D^2 = (D_C - D_{{}_0a(k+1)\,{}_0a'(k+1)})^{-1} \qquad (15.12.30)$$

$$ {}_0U = {}_0D^2 \; {}_0a(k+1) \qquad (15.12.31)$$

$$ {}_0W = C \; {}_0U - {}_0a(k+1) \qquad (15.12.32)$$

$$ {}_0t \; {}_0t' = {}_0W' \; {}_0U \qquad (15.12.33)$$

$$ {}_1a = {}_0W \; {}_0t'^{-1} \qquad (15.12.34)$$

and in general

$$ {}_iD^2 = (D_C - D_{{}_ia(k+1)\,{}_ia'(k+1)})^{-1} \qquad (15.12.35)$$

$$ {}_iU = {}_iD^2 \; {}_ia(k+1) \qquad (15.12.36)$$

$$ {}_iW = C \; {}_iU - {}_ia(k+1) \qquad (15.12.37)$$

$$ {}_it \; {}_it' = {}_iW' \; {}_iU \qquad (15.12.38)$$

$$ {}_{i+1}a = {}_iW \; {}_it'^{-1} \qquad (15.12.39)$$

We may continue until

$$ {}_it_{(k+1),\,(k+1)} < 0 \qquad (15.12.40)$$

or sooner.

Here also the solution is independent of scale, as shown in Section 15.12.7.

15.12.4 Proof of the Communality Successive Factor Vector Method

Let $\underline{C}$ be the covariance matrix and consider the rank reduction vector $\underline{{}_0a}$ given by

$${}_0a = C\, V\, (V'\, C\, V)^{-\frac{1}{2}} \qquad (15.12.41)$$

where $\underline{V}$ is arbitrary.

We indicate a diagonal of the elements of $\underline{{}_0a}$ by

$${}_0D = D^{-1}_{{}_0a} \qquad (15.12.42)$$

Let

$${}_1a = C\, {}_0D^2\, {}_0a\, ({}_0a'\, {}_0D^2\, C\, {}_0D^2\, {}_0a)^{-\frac{1}{2}} \qquad (15.12.43)$$

In general, let

$${}_iD = D^{-1}_{{}_ia} \qquad (15.12.44)$$

and

$${}_{i+1}a = C\, {}_iD^2\, {}_ia\, ({}_ia'\, {}_iD^2\, C\, {}_iD^2\, {}_ia)^{-\frac{1}{2}} \qquad (15.12.45)$$

Let

$${}_iU = D^{-1}_{{}_ia}\, 1 \qquad (15.12.46)$$

From Eqs. (15.12.44) and (15.12.46)

$${}_iU = {}_iD^2\, {}_ia \qquad (15.12.47)$$

From Eq. (15.12.47) in Eq. (15.12.45)

$$_{(i+1)}a = C\ {}_iU\ ({}_iU'\ C\ {}_iU)^{-\frac{1}{2}} \tag{15.12.48}$$

From Eqs. (15.12.44) and (15.12.47)

$$_iD = {}_iD_U \tag{15.12.49}$$

If we let

$$_iW = C\ {}_iU \tag{15.12.50}$$

we have as the computational sequence

$$D_{_iU} = D^{-1}_{_ia} \tag{15.12.51}$$

$$_iW = C\ {}_iU \tag{15.12.52}$$

$$\alpha_i = \frac{1}{\sqrt{{}_iW'\ {}_iU}} \tag{15.12.53}$$

$$_{i+1}a = {}_iW\ \alpha_i \tag{15.12.54}$$

We may continue Eqs. (15.12.51) through (15.12.54) until $\underline{\alpha_i}$ stabilizes, at which point

$$_ia = a_{.1} \tag{15.12.55}$$

That this solution is independent of scale can be readily seen by writing the general form from Eq. (15.12.45) as

$$a = C\ D_a^{-1}\ 1\ (1'\ D_a^{-1}\ C\ D_a^{-1}\ 1)^{-\frac{1}{2}} \tag{15.12.56}$$

Then consider any scaling diagonal $\underline{\Delta}$ such that

$$A = \Delta a \tag{15.12.57}$$

$$\gamma = \Delta C \Delta \tag{15.12.58}$$

and

$$A = \gamma D_A^{-1} 1 (1' D_A^{-1} C D_A^{-1} 1)^{-\frac{1}{2}} \tag{15.12.59}$$

Substituting Eqs. (15.12.57) and (15.12.58) in Eq. (15.12.59)

$$\Delta a = \Delta C \Delta (\Delta^{-1} D_a^{-1} 1) [1' D_a^{-1} \Delta^{-1} (\Delta D \Delta) \Delta^{-1} D_a^{-1} 1]^{-\frac{1}{2}} \tag{15.12.60}$$

Eq. (15.12.60) reduces at once to Eq. (15.12.56).

Once $\underline{a_{.1}}$ is obtained, we can solve for a residual matrix.

$${}_1C = C - a_{.1} a'_{.1} \tag{15.12.61}$$

and operate on $\underline{{}_1C}$ as before to obtain $\underline{a_{.2}}$. The procedure is readily generalized to any number of factors.

15.12.5 Proof of the Communality Factor Matrix Method

Let $\underline{C}$ be the covariance matrix and consider the approximate solution

$$C - a a' = {}_1C \tag{15.12.62}$$

where the width of $\underline{a}$ is chosen.

We let

$$D^{-2} = D_{a a'} \tag{15.12.63}$$

and

$$D a = \alpha \tag{15.12.64}$$

Consider also

$$D\ R\ D\ \alpha = \alpha\ (\alpha'\ \alpha) \qquad (15.12.65)$$

We assume $\underline{a}$ of fixed width and start with, say, a principal axis or some other approximation to $\underline{a}$. We call this solution $\underline{{}_0a}$. We let

$${}_0D = D^{-\frac{1}{2}}_{{}_0a\ {}_0a'} \qquad (15.12.66)$$

and

$${}_0\alpha = {}_0D\ {}_0a \qquad (15.12.67)$$

We then consider the matrix reduction solution

$${}_1a\ {}_1a' = C\ {}_0D\ {}_0\alpha\ ({}_0\alpha'\ {}_0D\ C\ {}_0D\ {}_0\alpha)^{-1}\ {}_0\alpha'\ {}_0D\ C \qquad (15.12.68)$$

We let

$${}_0U = {}_0D^2\ {}_0a \qquad (15.12.69)$$

From Eqs. (15.12.64), (15.12.68), and (15.12.69)

$${}_1a\ {}_1a' = R\ {}_0U\ ({}_0U'\ R_0\ {}_0U)^{-1}\ {}_0U'\ R \qquad (15.12.70)$$

We let

$${}_0W = R\ {}_0U \qquad (15.12.71)$$

$${}_0t\ {}_0t' = {}_0W'\ {}_0U \qquad (15.12.72)$$

$${}_1a = {}_0W\ {}_0t'^{-1} \qquad (15.12.73)$$

or in general

$${}_1D = D^{-\frac{1}{2}}_{{}_1a\ {}_1a'} \qquad (15.12.74)$$

$$ {}_iU = {}_iD^2 \, {}_ia \tag{15.12.75}$$

$$ {}_iW = C \, {}_iU \tag{15.12.76}$$

$$ {}_it \, {}_it' = {}_iW' \, {}_iU \tag{15.12.77}$$

$$ {}_{i+1}a = {}_iW \, {}_it'^{-1} \tag{15.12.78}$$

Eqs. (15.12.74) through (15.12.78) continue until in

$$ tr\,({}_{i+1}t) - tr\,({}_it) = P \tag{15.12.79}$$

$\underline{P}$ is sufficiently small. As $\underline{I}$ increases, $\underline{{}_it}$ will approach a diagonal matrix of the basic structure of $\underline{{}_iD \; R \; {}_iD}$ or, in general,

$$ D_{{}_it} \rightarrow {}_i\alpha' \, {}_i\alpha \tag{15.12.80}$$

We can show by Section 15.12.7 that this procedure is independent of scale.

15.12.6 Proof of the Communality Progressive Factor Matrix Method

Let Eqs. (15.12.74) through (15.12.78) be the iterative procedure for $\underline{a}$ of width $\underline{k}$ where $\underline{k} \geq 1$. We continue until Eq. (15.12.79) is satisfied and let

$$ a_{(k)} = (a_{.1} \ldots a_{.k}) \tag{15.12.81}$$

We then let

$$ a_{(k+1)} = (a_{(k)} \,,\, {}_0a) \tag{15.12.82}$$

where $\underline{{}_0a}$ is determined in some suitable manner. In particular, we may consider

$$ {}_{k+1}C = C - a_{(k)}\, a'_{(k)} \qquad (15.12.83) $$

and let

$$ {}_0a = {}_{k+1}C\, 1\, (1'\, {}_{k+1}C\, 1)^{-\frac{1}{2}} \qquad (15.12.84) $$

where obviously the operations in Eq. (15.12.84) can be performed directly with $\underline{C}$ and $\underline{a_{(k)}}$ so that $\underline{{}_{k+1}C}$ need not be computed.

We then let

$$ {}_0D^2 = D^{-1}_{{}_0a(k+1)\, {}_0a'(k+1)} \qquad (15.12.85) $$

$$ {}_0U = {}_0D^2\, {}_0a(k+1) \qquad (15.12.86) $$

$$ {}_0W = C\, {}_0U \qquad (15.12.87) $$

$$ {}_0t\, {}_0t' = {}_0W'\, {}_0U \qquad (15.12.88) $$

$$ {}_1a = {}_0a + {}_0W\, {}_0t'^{-1} \qquad (15.12.89) $$

and in general,

$$ {}_iD^2 = D^{-1}_{{}_ia(k+1)\, {}_ia'(k+1)} \qquad (15.12.90) $$

$$ {}_iU = {}_iD^2\, {}_ia(k+1) \qquad (15.12.91) $$

$$ {}_iW = C\, {}_iU \qquad (15.12.92) $$

$$ {}_it\, {}_it' = {}_iW'\, {}_iU \qquad (15.12.93) $$

$$ {}_{i+1}a = {}_ia\, {}_iW\, {}_it'^{-1} \qquad (15.12.94) $$

We continue until

$$\left| 1 - \frac{\operatorname{tr}(t_k)}{\operatorname{tr}(C)} \right| = P \tag{15.12.95}$$

is sufficiently small.

Here also the solution is independent of scale as shown in Section 15.12.7.

15.12.7 Proof of the Generalized Procedure Independent of Scale

Given the $\underline{n} \times \underline{n}$ covariance matrix $\underline{C}$ and the $\underline{n} \times \underline{k}$ factor matrix $\underline{a}$. Consider

$$E = C - a\,a' \tag{15.12.96}$$

Let

$$D = g\,D_C + f\,D_{a\,a'} \tag{15.12.97}$$

where $\underline{g}$ and $\underline{f}$ are scalars. Consider the basic structure rank reduction solution

$$a = C\,D^{-1}\,a\,(a'\,D^{-1}\,C\,D^{-1}\,a)^{-\frac{1}{2}}\,h \tag{15.12.98}$$

where $\underline{h}$ is a square orthonormal. Let $\underline{\Delta}$ be diagonal and

$$\gamma = \Delta\,C\,\Delta \tag{15.12.99}$$

$$\Lambda = \Delta\,a \tag{15.12.100}$$

$$\delta = g\,D_\gamma + f\,D_{AA'} \tag{15.12.101}$$

and consider the basic structure rank reduction solution

$$\Lambda = \gamma\, \delta^{-1} A\, (A'\, \delta^{-1} \gamma\, \delta^{-1} A)^{-\frac{1}{2}}\, h \qquad (15.12.102)$$

From Eqs. (15.12.96), (15.12.99), and (15.12.100) in (15.12.101)

$$\delta = \Delta^2 D \qquad (15.12.103)$$

From Eqs. (15.12.99), (15.12.100), and (15.12.103) in (15.12.102)

$$a = C\, D^{-1} a\, (a'\, D^{-1} C\, D^{-1} a)^{-\frac{1}{2}}\, h \qquad (15.12.104)$$

which is the same as Eq. (15.12.98).

Now we let

$$\alpha = \delta^{-\frac{1}{2}} A \qquad (15.12.105)$$

From Eqs. (15.12.100) and (15.12.103) in (15.12.105)

$$\alpha = D^{-\frac{1}{2}} a \qquad (15.12.106)$$

which shows that α is independent of Δ and depends only on g and f in Eq. (15.12.97). If we let $g = 0$, $f = 1$, we get the communality scaling type solutions. If we let $g = 1$, $f = 0$, we get the conventional basic structure solution applied to the correlation matrix.

We may now substitute $C - D$ for C, and $\gamma - \delta$ for γ, and show that α is independent of Δ and depends only on g and f. Then for $g = 1$ and $f = -1$ we get the specificity scaling type of solutions.

CHAPTER 16

IMAGE ANALYSIS

In earlier chapters we have referred to the communality problem and have indicated that from a computational point of view we may be concerned with solving for unknown diagonal elements of the correlation or covariance matrix in such a way that the modified matrix is of lower rank than the original. Presumably the original matrix will be basic in most cases.

Whether we can select diagonal elements so that the matrix without alterations in the offdiagonal elements is of lower rank than the original is a question of fact. We know that in some cases this cannot be done except for a rank reduction of 1. It is well known that experimental covariance matrices can in general be reduced to a rank one less than their order by a change in one diagonal element.

There are, of course, as many diagonal elements as the order of the matrix, so that there would be $\underline{n}$ ways that the matrix could be reduced in rank by at least 1. Actually, in the case of the correlation matrix, we know that this diagonal element which for any particular variable will reduce the rank of the matrix by 1, is the squared multiple correlation of that variable with all of the others.

The problem of communality and how to solve for the unknown elements has been very troublesome over the years. Many investigators are becoming convinced that the questions have not been properly stated and that the problem has not been properly formulated.

We have seen in the previous chapter that another way of looking at the problem is from the scaling point of view. It should be remembered that the communality problem has its origin and basic motivation from a consideration

of hypotheses about the kinds of factors which may exist in a set of variables. These are factors which are common to two or more of the variables, and factors which are specific to each of the variables. This does not, however, suggest specifically a mathematical formulation, because the unique variance must consist of both systematic specific variance, and error or unsystematic specific variance.

Another approach to the solution of the issues involved in the communality controversy has in recent years received considerable emphasis under the impetus of Louis Guttman (1953). His work has offered some hope of getting out of some of the dilemmas and contradictions involved in the traditional formulations of the communality problem. This approach is based on what he terms image analysis. The notion here is that a factor analysis should be concerned primarily with that part of each variable which can be estimated from all of the other variables in the set, and that as much as possible of the specific variance should be eliminated.

We shall therefore consider in this chapter a group of methods based on Guttman's image analysis which are somewhat different from those considered in the previous chapter. The methods of Chapter 15 operate upon a transformed score matrix in which the transformation consists of multiplication by a scaling diagonal. In the examples here, we also work with the transformed score matrix, in that the matrix is multiplied on the right by another matrix. The matrix in this case, however, is not a diagonal matrix but a more general type of matrix which we shall develop in detail in the following sections.

16.1 Characteristics of the Methods

We shall consider now the characteristics which are common to all of

these methods. First, all of the methods are based on the image matrix, which is the matrix consisting of the part of each variable which can be predicted by all the remaining ones. Second, the calculation of the inverse of the correlation matrix is required in all of the solutions. The problem of scaling of the variables is a consideration, but the methods differ essentially in the arbitrary scaling procedures adopted. Third, the correlation matrix should be basic. Finally, the solutions are, in general, basic structure solutions.

16.1.1 The Communality Score Matrix. The basis for all of these methods is a matrix derived from the data matrix by conventional least square procedures. In effect, one gets the best least square estimate of each attribute vector in the data matrix from all of the remaining $\underline{n}$-1 vectors. In this way one gets a matrix of least square estimates of the data matrix vectors. Actually, one does not go through the tedious and detailed operations of calculating the regression equations and the estimated vectors. By algebraic shortcuts one arrives at a matrix which transforms the original data matrix into this so-called image or estimated data matrix. Also, in practice, one does not operate directly on the data matrix, but rather on a correlation or covariance matrix derived from it.

16.1.2 Calculation of the Inverse. The methods considered in this chapter differ essentially from all of those we have considered previously in that the calculation of the inverse of the correlation matrix is required as a basis for the solution. For this reason the computations can be considerably more involved than previous methods we have considered. The computation of the inverse of a very large matrix involving several hundred or more variables is in itself an appreciable computational enterprise. It is

therefore only since the advent of the high speed electronic computers that methods of factor analysis based upon the image analysis approach of Guttman have become feasible.

16.1.3 Scaling Considerations. The essential differences in the methods we shall consider are those involving scaling of the variables after the original data matrix has been converted to an image or estimated matrix. These scaling methods are, however, considerably simpler than those in the previous chapter where, as we recall, elaborate iteration procedures were required to arrive at a scaling matrix for each of the models presented. In this set of methods, one adopts a simple and perhaps arbitrary rationale for the scaling of the variables, and, proceeding from this scaling on the covariance matrix of the estimated variables, one does not attempt to alter it by successive approximations as in the previous chapter. It is possible with these models to solve for scaling constants as in Chapter 15. However, the fruitfulness of such approaches has not yet been demonstrated.

16.1.4 The Basic Correlation Matrix. In the methods discussed in this chapter we must have a basic correlation matrix. This, of course, follows from the fact that we work with the inverse of the correlation or covariance matrix as part of the general procedure. In the past it has been true that most of the correlation matrices on which factor analyses were performed were basic and did have a regular inverse. Therefore this restriction in the method has not been a practical or serious one. However, we may very well encounter correlation matrices which are not basic. The most obvious case is the one in which we have more attributes than entities or persons. An example is the personality inventory for which we wish to consider each item on

the inventory as a variable. We may have many hundreds of items in the inventory, and it may be administered to only a hundred persons. Actually, in the experimental situation, one is faced with the problem that the longer the inventory to be administered for research purposes, the more difficult it is to amass cases or entities which have responded to the inventory. Therefore, in practical situations there is a tendency to find an inverse relationship between the number of variables and the number of entities. This is a consequence of the limited time which potential subjects have available for taking the inventory.

In any case, the correlation or covariance matrix cannot have rank greater than the number of entities. The problem of how the concept of the general inverse of the matrix could be used in connection with the image analysis type of factor analytic models has not yet been explored. Whether or not this would be a fruitful approach, even if it were mathematically and computationally feasible, requires further investigation.

16.1.5 Basic Structure Solutions. It is quite possible to apply any of the factor analytic procedures we have discussed in previous chapters to the types of transformed covariance or correlation matrices which we work with in the image analysis models. Even the centroid, the group centroid, and the multiple group methods can be applied. However, after going to all the trouble of calculating the inverse of the matrix, and considering the advantages of the basic structure solution, one would in general adopt some basic structure type of solution, particularly when high speed computers are available. The methods based on the image analysis approach would not be feasible for sizable sets of data matrices if only desk computers were available. As a matter of fact, it is only for small demonstrations or fictitious

examples that one would be likely to use desk computers for the types of models outlined below.

16.2 Kinds of Methods

As indicated above, all of the methods start with a covariance matrix which consists of the variances and covariances of the least square estimated variables. We assume that the data matrix has been reduced to standard or normalized form. This has been done before the symbolic transformation to the estimated or image variables has been accomplished.

On the basis of this assumption, we then have four different variations of the factoring procedure. First, the estimated covariance matrix may be factored. Second, the correlation matrix of the estimated variables may be factored. Third, the covariance matrix of estimated variables may be scaled in such a way that it is independent of the scaling of the original variables. Finally, the inverse of the covariance matrix may be scaled so that it yields the best least square approximation to the identity matrix.

16.2.1 The Estimated Covariance Matrix. As indicated above, the matrix of variances and covariances of the variables estimated by least square regression from the data matrix of normalized variables can be obtained by suitable mathematical transformation of the correlation matrix. After this transformation has been applied to the correlation matrix, the resulting covariance matrix is subjected to a basic structure type solution such as an eigenvalue-eigenvector solution in which the largest roots and corresponding vectors are extracted without further alteration of this covariance matrix.

16.2.2 The Estimated Correlation Matrix. Instead of working with the covariance matrix of estimated variables, one may wish to work with the actual

correlation matrix of these estimated variables. This is a simple matter, for one can merely pre- and postmultiply the estimated covariance matrix by the reciprocal square root of its diagonal, or what amounts to the same thing, by the reciprocal square root of the variances of the estimated variables. This correlation matrix with unity in the diagonal is now factored in the conventional manner by basic structure or eigenvalue procedures.

16.2.3 The Independent Scale Procedure. It will be recognized that both of the variations considered above are arbitrary from the scaling point of view. One may, however, prefer a model which does not depend on the assumption of either a standardized data or an image matrix, but a method which is independent of the scaling of the original or image variables. In other words, one may wish to use a procedure so that, for any scaling diagonal one may use on the data matrix, this diagonal cancels out in the covariance matrix which is finally adopted for factoring. The third method achieves this objective.

16.2.4 The Optimal Residual Matrix. In the image analysis approach, each vector of the image matrix is defined as the part of each variable which can be predicted from all of the others. Implicit also is the concept of the anti-image matrix which consists of that part of each variable which cannot be predicted from any of the others. We may therefore also define an anti-image covariance matrix as indicated in Section 16.7. This method proceeds on the assumption that the image covariance matrix should be scaled in such a way that when the anti-image covariance matrix is scaled in the same way, it will be as close to an identity matrix as possible. In other words, the rationale is that the anti-image covariance matrix shall have a scaling such that, compared to the variances, the covariances will be as small as possible

in the least square sense.

16.3 The Image Covariance Matrix

16.3.1 Characteristics of the Method. In this method, as indicated above, the analysis is performed directly on the covariance matrix of the variables which have been estimated from the normalized data matrix by the least square model. It can be seen from the mathematical proof of this method in Section 16.7 that the diagonal or variances of this covariance matrix consist of the squared multiple correlations of each variable with all of the others.

In general, any method of factor analysis based on basic structure procedures or approximations to them tend to give the greatest weight to the variables with the largest variances. That is, if a covariance rather than a correlation matrix is operated upon, other things being equal, the variables with the largest variances have the greatest weight or influence in determining the factor loadings for that matrix. It can be seen, therefore, that since the variances are squared multiple correlation coefficients, those variables which have the highest multiple correlation with all of the other variables receive the greatest weight in the determination of the basic structure factor matrix. Conversely, variables with very low multiple correlations receive very little weight in the solution. In particular, if a variable is entirely independent of the others--that is, if it has zero correlations--it receives no weight whatever, and will therefore not have a loading in any of the factors.

The rationale here can obviously be defended if one takes the position that he is interested only in the factor loadings for those factors which are common to two or more of the variables. This is the traditional Thurstonian

approach to the problem. The communality concept appears to make more sense from the mathematical and theoretical point of view via the image analysis approach than via the more traditional approach in which one alters the diagonal elements without altering the correlation coefficients. It would seem that any defensible approach should be based on some transformation of the data matrix, rather than on the data matrix plus something which is not connected in any way with the data matrix. This latter is implicit in the conventional communality approach.

16.3.2 Computational Equations

16.3.2a Definition of Notation

$\underline{R}$ is the correlation matrix.

$\underline{R}^{-1}$ is the inverse of $\underline{R}$.

$\underline{D_\rho}$ is the diagonal of $\underline{\rho}$.

$\underline{C_{W\,W}}$ is the estimated covariance matrix.

$\underline{a}$ is the factor loading matrix.

16.3.2b The Equations

$$\rho = R^{-1} \tag{16.3.1}$$

$$D = D_\rho^{-1} \tag{16.3.2}$$

$$C^\rho = \rho - D_\rho \tag{16.3.3}$$

$$r^\rho = D\, C^\rho\, D \tag{16.3.4}$$

$$_{C}{}^{R} = R - D \tag{16.3.5}$$

$$C_{W\,W} = {}_{C}{}^{R} + {}_{r}{}^{\rho} \tag{16.3.6}$$

$$C_{W\,W} = Q\,\delta\,Q' \tag{16.3.7}$$

$$a = Q\,\delta^{\frac{1}{2}} \tag{16.3.8}$$

16.3.3 Computational Instructions. We assume that the matrix of correlations of the variables to be factored is available. This matrix we indicate by R. The first step is to calculate its inverse by one of the conventional procedures. Eq. (16.3.1) indicates the inverse of the correlation matrix. We designate this by ρ.

Eq. (16.3.2) is the diagonal matrix whose elements are the reciprocals of the diagonal elements in the matrix given by Eq. (16.3.1).

We next indicate the inverse of the correlation matrix with the diagonal elements removed or made 0, as in Eq. (16.3.3).

The next step is the pre- and postmultiplication of the matrix given by Eq. (16.3.3), by the diagonal matrix of Eq. (16.3.2). This is indicated by Eq. (16.3.4).

The next step consists of subtracting from the original correlation matrix the diagonal matrix of Eq. (16.3.2). This is indicated in Eq. (16.3.5). It can be shown that the diagonal elements in the matrix on the left hand side of Eq. (16.3.5) are now the squared multiple correlation coefficients of each variable with the remaining variables.

We next add together the matrices given by Eqs. (16.3.4) and (16.3.5), as indicated in Eq. (16.3.6). This is now the covariance matrix of the estimated variables, or the image covariance matrix.

Eq. (16.3.7) indicates the basic structure solution of this image covariance matrix.

The analysis may be carried to as many factors as desired. There seems to be no very good rule for this particular model, but a rough rule-of-thumb criterion is that the sum of the currently calculated basic diagonals be approximately 80 to 85 percent of the trace of the image covariance matrix given by Eq. (16.3.6).

Eq. (16.3.8) gives the factor loading matrix as a function of the basic diagonal and the basic orthonormal vectors indicated in Eq. (16.3.7).

16.3.4 Numerical Example. The correlation matrix used here is the same as in previous chapters.

Table 16.3.1 gives the image covariance matrix of the correlation matrix. The inverse of the correlation matrix is not displayed, although it can be printed out from the appropriate Fortran program if desired. The procedure for calculating the inverse is given in Chapter 3, Section 3.5. This is included as "Subroutine Symin" in the Fortran listing.

The first row of Table 16.3.2 gives all the basic diagonal elements of the image covariance matrix. The body of the table gives the column vectors of the left basic orthonormal. These are, of course, proportional to the corresponding factor loading vectors of the image covariance matrix. The basic orthonormal matrix must be postmultiplied by the square root of the basic diagonal matrix to yield the principal axis vectors. These have not been calculated. The basic structure factors are probably of more interest than the principal axis vectors, although the latter could readily be obtained by several additional statements in the Fortran program.

Table 16.3.1 - Image Covariance Matrix

0.7331	0.6880	0.6621	0.1204	0.0652	0.1011	0.2936	0.3018	0.3356
0.6880	0.7444	0.6659	0.1472	0.0672	0.1167	0.3163	0.3067	0.3577
0.6621	0.6659	0.6905	0.2146	0.1693	0.2389	0.2877	0.3047	0.3554
0.1204	0.1472	0.2146	0.5114	0.4875	0.4942	0.2182	0.1784	0.2853
0.0652	0.0672	0.1693	0.4875	0.5720	0.4858	0.1611	0.0814	0.2464
0.1011	0.1167	0.2389	0.4942	0.4858	0.5624	0.1602	0.1417	0.2625
0.2936	0.3163	0.2877	0.2182	0.1611	0.1602	0.4808	0.3964	0.4177
0.3018	0.3067	0.3047	0.1784	0.0814	0.1417	0.3964	0.4989	0.3885
0.3356	0.3577	0.3554	0.2853	0.2464	0.2625	0.4177	0.3885	0.4359

Table 16.3.2 - Basic Diagonal and Basic Orthonormal of Image Covariance Matrix

3.0137	1.3337	0.5792	0.1254	0.0651	0.0508	0.0371	0.0214	0.0029

-0.4108	0.3458	-0.2489	0.0805	-0.4535	0.0380	0.6019	0.2427	-0.1224
-0.4221	0.3318	-0.2246	0.1425	0.4832	-0.4261	-0.3195	0.1327	-0.3261
-0.4310	0.2016	-0.3089	-0.2148	-0.0538	0.3573	-0.2874	-0.4966	0.4149
-0.2587	-0.4596	-0.0780	-0.0329	0.1784	-0.5532	0.3880	-0.1702	0.4442
-0.2172	-0.5232	-0.1902	0.4109	-0.5056	-0.0719	-0.3804	-0.0709	-0.2522
-0.2472	-0.4862	-0.1926	-0.4374	0.3164	0.4173	0.1472	0.1934	-0.3752
-0.3027	-0.0133	0.5306	0.4669	0.2479	0.2779	0.2532	-0.4101	-0.1940
-0.2924	0.0468	0.5653	-0.5696	-0.3283	-0.2947	-0.1488	-0.0951	-0.2063
-0.3391	-0.0765	0.3343	0.1533	0.0352	0.1999	-0.2234	0.6544	0.4739

16.4 The Image Correlation Matrix

16.4.1 Characteristics of the Method. The method is similar to most of the conventional methods of factor analysis we have considered in previous chapters in that we begin with the matrix of the correlations of measures with unity in the diagonal. Here we make an arbitrary assumption that, for the particular sample, each of the estimated variables should have equal variance. The rationale or justification for this assumption is probably no worse or better than such an assumption for the original data matrix. If, however, one assumes that variables which correlate low with others should not, therefore, be weighted less in the factor solution, the procedure of using unit variances for the image variables is justified. In any case, this method does give relatively more weight to the variables which have the greater unique variance than does the previous method.

16.4.2 Computational Equations

16.4.2a Definition of Notation

$\underline{C}_{W\,W}$ is the estimated covariance matrix.

$D_{R^{-1}}$ is the diagonal of $\underline{R}^{-1}$.

$\underline{R}_{W\,W}$ is the estimated correlation matrix.

$\underline{Q\,\delta\,Q'}$ is the basic structure of $\underline{R}_{W\,W}$.

16.4.2b The Equations

Given the $\underline{C}_{W\,W}$ matrix

$$d = (I - D_{R^{-1}})^{\frac{1}{2}} \qquad (16.4.1)$$

$$R_{WW} = d^{-1} C_{WW} d^{-1} \tag{16.4.2}$$

$$R_{WW} - Q \delta Q' = \epsilon \tag{16.4.3}$$

$$a = Q \delta^{\frac{1}{2}} \tag{16.4.4}$$

16.4.3 Computational Instructions. The computational instructions for this procedure involve only a few more steps than those of the previous method. We begin in the same way by calculating the covariance matrix of the image variables, that is, the $\underline{C_{WW}}$ matrix. We have seen that the diagonal elements of this covariance matrix are given by the identity less a diagonal matrix which is the inverse of the diagonal elements of the inverse of the correlation matrix. This is indicated in Eq. (16.4.1). It is precisely a diagonal matrix of the squared multiple correlation coefficients of each variable with all the remaining variables.

The next set of computational steps is given in Eq. (16.4.2). This consists of pre- and postmultiplying the covariance image matrix by the inverse of the $\underline{d}$ matrix calculated in Eq. (16.4.1). The resulting matrix is the correlation matrix of the image variables.

The next set of computations consists in finding the basic structure factor vectors for the required number of factors, as indicated in Eq. (16.4.3).

The factor loading matrix is indicated in the conventional manner in Eq. (16.4.4).

16.4.4 Numerical Example. We use the same correlation matrix as in the previous section. Here we begin with the image covariance matrix solved for in the previous section.

Table 16.4.1 gives the correlation matrix obtained from the image covariance matrix. This is obtained by pre- and postmultiplying the image covariance matrix by the reciprocal square root of its diagonal.

The first row of Table 16.4.2 gives the elements of the basic diagonal of the image correlation matrix. The body of the table gives the left basic orthonormal of this matrix. It may be transformed to a factor loading matrix by the usual method indicated in Eq. (16.4.4).

16.5 The Independent Scale Matrix

16.5.1 Characteristics of the Method. The previous two methods which we considered were based on arbitrary scaling procedures. In the first case we required that the original variables be in standard score form, and in the second case we required that the image variables be in standard score form. It may be desirable to have a method which does not impose any such arbitrary scaling.

We therefore consider a method which scales the image covariance matrix in such a way as to cancel out any particular scaling which has been applied to the original or image variables.

This method also has some interesting characteristics which are indicated in Section 16.7.3. The particular scaling applied to the data matrix is such that the image covariance matrix is the sum of the covariance matrix of the scaled data matrix and the inverse of this covariance matrix, less twice the identity matrix.

16.5.2 Computational Equations

16.5.2a Definition of Notation

Table 16.4.1 - Image Correlation Matrix

1.0000	0.9314	0.9306	0.1966	0.1008	0.1574	0.4946	0.4990	0.5936
0.9314	1.0000	0.9289	0.2387	0.1029	0.1803	0.5287	0.5033	0.6279
0.9306	0.9289	1.0000	0.3612	0.2695	0.3834	0.4992	0.5192	0.6479
0.1966	0.2387	0.3612	1.0000	0.9015	0.9214	0.4400	0.3532	0.6042
0.1008	0.1029	0.2695	0.9015	1.0000	0.8566	0.3072	0.1524	0.4934
0.1574	0.1803	0.3834	0.9214	0.8566	1.0000	0.3081	0.2674	0.5301
0.4946	0.5287	0.4992	0.4400	0.3072	0.3081	1.0000	0.8093	0.9125
0.4990	0.5033	0.5192	0.3532	0.1524	0.2674	0.8093	1.0000	0.8330
0.5936	0.6279	0.6479	0.6042	0.4934	0.5301	0.9125	0.8330	1.0000

Table 16.4.2 - Basic Diagonal and Basic Orthonormal Matrix of Image Correlation Matrix

5.1559	2.2350	1.0679	0.2392	0.1099	0.0883	0.0589	0.0397	0.0052

-0.3300	-0.3557	-0.3292	-0.0630	-0.3118	0.0426	-0.6088	-0.4021	-0.1445
-0.3392	-0.3428	-0.3068	-0.1185	0.2553	-0.4410	0.5042	-0.0200	-0.3807
-0.3668	-0.2347	-0.3792	0.1657	0.0189	0.3361	0.0650	0.5459	0.4728
-0.3094	0.4512	-0.0757	0.0589	0.0797	-0.6697	-0.2413	-0.0117	0.4223
-0.2523	0.5038	-0.1594	-0.3643	-0.6019	0.1279	0.2244	0.1799	-0.2512
-0.2814	0.4648	-0.1780	0.3935	0.4708	0.3568	-0.1045	-0.1302	-0.3760
-0.3535	-0.0798	0.4927	-0.4978	0.3183	0.0554	-0.3357	0.3636	-0.1699
-0.3312	-0.1450	0.5081	0.6234	-0.3792	-0.1391	0.0655	0.1355	-0.1914
-0.4104	-0.0036	0.2999	-0.1644	0.0419	0.2760	0.3666	-0.5830	0.4030

$\underline{C_{WW}}$ is the estimated covariance matrix.

$\underline{G}$ is the estimated covariance matrix independent of scale.

$\underline{Q\,\delta\,Q'}$ is the basic structure of $\underline{G}$.

16.5.2b The Equations

Given the $\underline{C_{WW}}$ matrix

$$d = D_{R^{-1}}^{\frac{1}{2}} \tag{16.5.1}$$

$$G = d\,C_{WW}\,d \tag{16.5.2}$$

$$G = Q\,\delta\,Q' \tag{16.5.3}$$

$$a = Q\,\delta^{\frac{1}{2}} \tag{16.5.4}$$

16.5.3 Computational Instructions. This method, like the previous one, begins with the covariance matrix of the image variables. It may or may not be based on an image covariance matrix derived from standard measures. The result is the same whether it is applied to a matrix derived from standard measures or to a matrix derived from arbitrarily scaled measures. For convenience we shall assume that the image covariance matrix is based on standardized measures.

We begin with a diagonal matrix, as indicated in Eq. (16.5.1). This is a diagonal matrix made up of the square roots of the diagonal elements in the inverse of the correlation matrix.

We now pre- and postmultiply the covariance matrix of the image variables by this diagonal matrix, as indicated in Eq. (16.5.2). This we call

the $\underline{Q}$ matrix. This matrix now has the interesting property that it is the sum of a matrix and its inverse less twice the identity matrix.

Eq. (16.5.3) indicates the basic structure resolution of the $\underline{Q}$ matrix.

Eq. (16.5.4) gives the principal axis factor loading vectors for the specified number of factors.

16.5.4 Numerical Example. We use the correlation matrix as in the previous sections and begin with the image covariance matrix as calculated in those examples.

The first row in Table 16.5.1 gives the basic diagonal elements of the scale-free image covariance matrix. The body of the table gives the left basic orthonormal of this matrix. Perhaps the most striking feature of this table, as compared with corresponding tables for the two preceding methods, is the large first eigenvalue of 8.9569.

16.6 The Optimal Residual or Anti-Image Matrix

16.6.1 Characteristics of the Method. This method is somewhat different in rationale from the previous methods. It begins, as they do, with a covariance matrix of the image variables, but the rationale for the scaling procedure is less arbitrary than in the first two, although perhaps in a sense more arbitrary than for the third method.

Here the essential consideration is one developed independently by Guttman (1956) and Harris (1962). They were concerned with a scaling rationale for which the scaled anti-image covariance matrix yields the best least square approximation to an identity matrix. The method is of particular interest because of the recent work of Harris (1962) in which he has been concerned with the estimation of diagonal elements in the correlation matrix which will yield the best approximation to a lower rank approximation. This is the conventional communality problem.

Table 16.5.1 - Basic Diagonal and Basic Orthonormal of Scale-Free Image Matrix

8.9569	3.2956	1.2383	0.2787	0.2096	0.1196	0.1001	0.0466	0.0066

-0.5144	0.2675	-0.1628	-0.0754	0.6286	0.1830	-0.4206	0.1335	-0.0905
-0.5353	0.2518	-0.1301	-0.2229	-0.6625	0.2269	0.1128	0.1484	-0.2414
-0.4803	0.0791	-0.2376	0.3257	0.0569	-0.4390	0.3673	-0.3991	0.3322
-0.1609	-0.4786	-0.0668	-0.0277	-0.1704	0.4936	-0.3349	-0.3726	0.4650
-0.1329	-0.5571	-0.2088	-0.4379	0.3024	0.0869	0.5231	0.0373	-0.2501
-0.1608	-0.5268	-0.1949	0.4610	-0.1566	-0.3049	-0.3482	0.2665	-0.3692
-0.2160	-0.1096	0.5587	-0.4044	-0.0401	-0.4185	-0.2722	-0.4061	-0.2233
-0.2166	-0.0592	0.6101	0.5023	0.1264	0.4083	0.3018	-0.0429	-0.2225
-0.2348	-0.1624	0.3601	-0.1329	-0.0104	-0.1904	0.0567	0.6503	0.5556

16.6.2 Computational Equations

16.6.2a Definition of Notation

$\underline{C_{WW}}$ is the estimated covariance matrix.

$\underline{\rho^{(2)}}$ is a matrix whose elements are the squares of the elements in $\underline{\rho}$.

$\underline{\gamma}$ is the covariance matrix with optimal residual variance components.

16.6.2b The Equations

$$\rho = R^{-1} \qquad (16.6.1)$$

$$P = \rho^{(2)} \qquad (16.6.2)$$

$$d\,1 = P^{-1}\,D_\rho\,1 \qquad (16.6.3)$$

$$D = d^{\frac{1}{2}}\,D_\rho \qquad (16.6.4)$$

$$\gamma = D\,C_{WW}\,D \qquad (16.6.5)$$

$$\gamma = Q\,\delta\,Q' \qquad (16.6.6)$$

$$a = Q\,\delta^{\frac{1}{2}} \qquad (16.6.7)$$

16.6.3 Computational Instructions. The computations begin with the image covariance matrix calculated as in the previous methods. However, we must go back now to a solution of the scaling diagonal. We begin with the inverse of the correlation matrix, as indicated in Eq. (16.6.1).

Next we square each of the elements of the inverse calculated in Eq. (16.6.1), as indicated in Eq. (16.6.2). The superscript 2 enclosed in parentheses means that each element of the matrix on the right hand side of the equation has been squared.

Next we calculate a vector, as indicated in Eq. (16.6.3). The right hand side of this equation shows that the vector consists of the diagonal elements of the inverse of the correlation matrix given by Eq. (16.6.1). We then premultiply this vector by the inverse of the matrix calculated in Eq. (16.6.2).

We now define a new diagonal matrix, as in Eq. (16.6.4). This matrix is obtained by taking the diagonal elements of the inverse of the correlation matrix and multiplying these by the square roots of the elements calculated in Eq. (16.6.3). It should be observed that the solution given by Eq. (16.6.3) does not indicate offhand that all elements in the $\underline{d}$ matrix must be positive. If they are not positive, of course, we cannot have real numbers for their square roots. This is a limitation of the method. Research to date seems to indicate that with most experimental data matrices, Eq. (16.6.3) will give all positive elements.

We next pre- and postmultiply the image covariance matrix by the diagonal matrix of Eq. (16.6.4) to get a $\underline{\gamma}$ matrix, as in Eq. (16.6.5). This is the matrix which we now factor.

The basic structure of this matrix is indicated in Eq. (16.6.6).

The factor loading matrix is given in Eq. (16.6.7). This is simply the usual principal axis factors calculated for the desired number of factors.

16.6.4 Numerical Example. We use the correlation matrix as in the previous section.

Table 16.6.1 gives the inverse of the correlation matrix.

The body of Table 16.6.2 is the inverse of the matrix obtained by squaring the elements of Table 16.6.1. The row at the bottom of the table contains the elements of the scaling diagonal.

The body of Table 16.6.3 gives the image covariance matrix scaled so that the corresponding anti-image matrix is the best least square approximation to the identity matrix.

The first row of Table 16.6.4 consists of the elements of the basic diagonal of the matrix in Table 16.6.3. The body of the table is the corresponding left basic orthonormal.

16.7 Mathematical Proofs

16.7.1 The Estimated Covariance Matrix

Given the data matrix $\underline{x}$ in standard measures. We consider another matrix, $\underline{W}$, such that each vector $\underline{W}_{.j}$ of $\underline{W}$ is the least square estimate of $\underline{x}_{.j}$ calculated from the remaining $\underline{n}$-1 vectors of $\underline{x}$. We let

$$R = \frac{x' \, x}{N} \qquad (16.7.1)$$

We let $\underline{\beta}$ be the matrix of regression coefficients for estimating each variable from the remaining $\underline{n}$-1 variables so that

$$W = x \beta \qquad (16.7.2)$$

It is well known that $\underline{\beta}$ is given by

$$\beta = I - R^{-1} D_{R^{-1}}^{-1} \qquad (16.7.3)$$

if the variables are standard measures and $\underline{D_{R^{-1}}}$ is the diagonal of $\underline{R^{-1}}$. We

Table 16.6.1 - Inverse of Correlation Matrix or $\underline{p}$

3.7463	-2.0659	-1.2821	0.0950	0.2822	-0.0594	-0.0314	-0.0542	-0.1026
-2.0659	3.9116	-1.3784	0.2582	0.0563	-0.0744	-0.0506	-0.3145	-0.0787
-1.2821	-1.3784	3.2313	-0.3794	-0.2692	0.0068	-0.0519	0.2176	-0.1694
0.0950	0.2582	-0.3794	2.0465	-0.7099	-0.6166	-0.1215	-0.0187	-0.3038
0.2822	0.0563	-0.2692	-0.7099	2.3365	-1.1916	0.1040	-0.0447	-0.0316
-0.0594	-0.0744	0.0068	-0.6166	-1.1916	2.2853	-0.1310	0.1764	-0.1155
-0.0314	-0.0506	-0.0519	-0.1215	0.1040	-0.1310	1.9261	-0.9902	-0.3731
-0.0542	-0.3145	0.2176	-0.0187	-0.0447	0.1764	-0.9902	1.9958	-0.5397
-0.1026	-0.0787	-0.1694	-0.3038	-0.0316	-0.1155	-0.3731	-0.5397	1.7726

Table 16.6.2 - Inverse of $\underline{P}$ Matrix, and Scaling Diagonals for Optimal Error or Anti-Image Matrix

0.0786	-0.0209	-0.0086	0.0006	-0.0011	0.0002	-0.0002	0.0006	-0.0002
-0.0209	0.0724	-0.0099	-0.0008	0.0005	-0.0001	0.0004	-0.0018	0.0002
-0.0086	-0.0099	0.0990	-0.0031	-0.0010	0.0005	0.0002	-0.0009	-0.0007
0.0006	-0.0008	-0.0031	0.2425	-0.0191	-0.0124	-0.0008	0.0009	-0.0071
-0.0011	0.0005	-0.0010	-0.0191	0.1987	-0.0526	-0.0004	0.0004	0.0007
0.0002	-0.0001	0.0005	-0.0124	-0.0526	0.2067	-0.0003	-0.0015	-0.0003
-0.0002	0.0004	0.0002	-0.0008	-0.0004	-0.0003	0.2884	-0.0706	-0.0062
0.0006	-0.0018	-0.0009	0.0009	0.0004	-0.0015	-0.0706	0.2701	-0.0219
-0.0002	0.0002	-0.0007	-0.0071	0.0007	-0.0003	-0.0062	-0.0219	0.3208
1.6103	1.6127	1.5790	1.2936	1.2807	1.2964	1.2215	1.1905	1.2497

Table 16.6.3 - Image Covariance Matrix Scaled for Optimal Anti-Image Matrix

1.9008	1.7867	1.6834	0.2508	0.1346	0.2110	0.5776	0.5785	0.6752
1.7867	1.9358	1.6957	0.3072	0.1387	0.2439	0.6230	0.5889	0.7208
1.6834	1.6957	1.7216	0.4384	0.3424	0.4891	0.5548	0.5728	0.7013
0.2508	0.3072	0.4384	0.8558	0.8077	0.8287	0.3447	0.2748	0.4611
0.1346	0.1387	0.3424	0.8077	0.9382	0.8066	0.2520	0.1241	0.3943
0.2110	0.2439	0.4891	0.8287	0.8066	0.9452	0.2538	0.2186	0.4252
0.5776	0.6230	0.5548	0.3447	0.2520	0.2538	0.7174	0.5764	0.6376
0.5785	0.5889	0.5728	0.2748	0.1241	0.2186	0.5764	0.7071	0.5779
0.6752	0.7208	0.7013	0.4611	0.3943	0.4252	0.6376	0.5779	0.6807

Table 16.6.4 - Basic Diagonal and Orthonormal for Scaled Image Covariance Matrix

6.4745	2.4373	0.9357	0.2035	0.1438	0.0915	0.0730	0.0380	0.0052

-0.4973	-0.2864	0.1653	0.1004	0.6053	-0.2447	0.4182	0.1503	-0.0974
-0.5075	-0.2653	0.1303	0.2162	-0.6671	-0.2084	-0.1616	0.1532	-0.2651
-0.4937	-0.1097	0.2564	-0.3168	0.0677	0.4550	-0.2944	-0.4056	0.3395
-0.1804	0.4946	0.0833	0.0303	-0.1885	-0.5371	0.2678	-0.3289	0.4594
-0.1408	0.5336	0.2088	0.4374	0.3220	-0.0198	-0.5343	0.0142	-0.2659
-0.1720	0.5154	0.2087	-0.4509	-0.1466	0.2876	0.3715	0.2610	-0.3842
-0.2281	0.1037	-0.5631	0.3897	-0.0515	0.3584	0.3170	-0.4330	-0.2217
-0.2181	0.0494	-0.5767	-0.5284	0.1431	-0.3830	-0.3355	-0.0583	-0.2346
-0.2643	0.1655	-0.3822	0.1272	-0.0054	0.2073	-0.0633	0.6497	0.5210

therefore have

$$W = x\,(I - R^{-1} D_{R^{-1}})$$ (16.7.4)

The matrix of residuals $\underline{E}$ is then, from Eq. (16.7.4)

$$E = x - W$$ (16.7.5)

From Eqs. (16.7.4) and (16.7.5) we have

$$E = x\,R^{-1} D_{R^{-1}}$$ (16.7.6)

We shall now consider the covariance matrices involving $\underline{x}$, $\underline{W}$, and $\underline{E}$. We let

$$C_{x\,W} = \frac{x'\,W}{N}$$ (16.7.7)

$$C_{x\,E} = \frac{x'\,E}{N}$$ (16.7.8)

$$C_{W\,W} = \frac{W'\,W}{N}$$ (16.7.9)

$$C_{W\,E} = \frac{W'\,E}{N}$$ (16.7.10)

$$C_{E\,E} = \frac{E'\,E}{N}$$ (16.7.11)

These are the formulas developed by Guttman (1953) and discussed by Harris (1962).

From Eqs. (16.7.1), (16.7.4), and (16.7.7)

$$C_{x\,W} = R - D_{R^{-1}}$$ (16.7.12)

From Eqs. (16.7.1), (16.7.5), and (16.7.8)

$$C_{x\,E} = D_{R^{-1}}$$ (16.7.13)

From Eqs. (16.7.1), (16.7.4), and (16.7.9)

$$C_{W\,W} = (I - D_{R^{-1}}^{-1} R^{-1})\, R\, (I - R^{-1} D_{R^{-1}}^{-1}) \tag{16.7.14}$$

or

$$C_{W\,W} = R - 2\, D_{R^{-1}}^{-1} + D_{R^{-1}}^{-1} R^{-1} D_{R^{-1}}^{-1} \tag{16.7.15}$$

From Eqs. (16.7.1), (16.7.4), (16.7.6), and (16.7.10)

$$C_{W\,E} = D_{R^{-1}}^{-1} - D_{R^{-1}}^{-1} R^{-1} D_{R^{-1}}^{-1} \tag{16.7.16}$$

From Eqs. (16.7.1), (16.7.6), and (16.7.11)

$$C_{E\,E} = D_{R^{-1}}^{-1} R^{-1} D_{R^{-1}}^{-1} \tag{16.7.17}$$

The relationships among the various covariance matrices are obvious and have been discussed by Guttman (1953) and Harris (1962).

One may now regard the covariance matrix of the "estimated" variables as the logical matrix to factor, since presumably it has removed from each variable that part which does not overlap with the other variables. We therefore let the basic structure of Eq. (16.7.15) be

$$C_{W\,W} = Q\, \delta\, Q' \tag{16.7.18}$$

and factor to the desired number of roots and vectors.

16.7.2 The Estimated Correlation Matrix

Suppose we do not wish to factor the covariance matrix of the estimated variables, but rather the actual correlation matrix of this estimated matrix. We have from Section 16.7.1

$$C_{W\,W} = R - 2\,D_R^{-1} + D_R^{-1} R^{-1} D_R^{-1} \qquad (16.7.19)$$

We let

$$d = D_{C_{W\,W}} \qquad (16.7.20)$$

The matrix we wish to factor is

$$R_{W\,W} = d^{-\frac{1}{2}}\,C_{W\,W}\,d^{-\frac{1}{2}} \qquad (16.7.21)$$

From Eqs. (16.7.19) and (16.7.20)

$$d = I - D_R^{-1} \qquad (16.7.22)$$

It is well known that Eq. (16.7.22) is a diagonal of the squared multiple correlations between each variable and the remaining n-1 variables of the set. We repeat the covariance matrix of the estimated variables with the observed variables.

$$C_{x\,W} = R - D_R^{-1} \qquad (16.7.23)$$

It is clear from Eqs. (16.7.22) and (16.7.23) that the diagonals of $C_{W\,W}$ and $C_{x\,W}$ are the same.

16.7.3 The Solution Independent of Scale

Consider again the covariance matrix of the estimated variables x, namely,

$$C_{W\,W} = R - 2\,D_R^{-1} + D_R^{-1} R^{-1} D_R^{-1} \qquad (16.7.24)$$

It may be desirable to consider a factoring of, say,

$$G = D \, C_{W W} \, D \qquad (16.7.25)$$

where G is independent of the scaling of the variables. That is, we remove the assumption of standardized measures in x. Suppose we let

$$C = \frac{x' \, x}{N} \qquad (16.7.26)$$

where now we place no restrictions on the scaling of x. We then rewrite Eq. (16.7.24) as

$$C_{W W} = C - 2 \, D_{C^{-1}}^{-1} + D_{C^{-1}}^{-1} \, C^{-1} \, D_{C^{-1}}^{-1} \qquad (16.7.27)$$

We let

$$C = d \, R \, d \qquad (16.7.28)$$

where d is an arbitrary scaling diagonal.

Suppose we let D in Eq. (16.7.25) be

$$D = D_{R^{-1}}^{\frac{1}{2}} \qquad (16.7.29)$$

From Eqs. (16.7.24), (16.7.25) and (16.7.29) we have

$$G = D_{R^{-1}}^{\frac{1}{2}} \, R \, D_{R^{-1}}^{\frac{1}{2}} - 2 \, I + D_{R^{-1}}^{-\frac{1}{2}} \, R^{-1} \, D_{R^{-1}}^{-\frac{1}{2}} \qquad (16.7.30)$$

Suppose now we write Eq. (16.7.24) in the form of Eq. (16.7.30)

$$G = D_{C^{-1}}^{\frac{1}{2}} \, C \, D_{C^{-1}}^{\frac{1}{2}} - 2 \, I + D_{C^{-1}}^{-\frac{1}{2}} \, C^{-1} \, D_{C^{-1}}^{-\frac{1}{2}} \qquad (16.7.31)$$

From Eqs. (16.7.28) and (16.7.31) we see that the d matrix cancels out, and Eq. (16.7.31) becomes precisely Eq. (16.7.30). Therefore G in Eq. (16.7.31)

is independent of scale and would seem to be a desirable matrix for factoring. Furthermore, suppose we let

$$g = D_R^{\frac{1}{2}-1} R \, D_R^{\frac{1}{2}-1} \tag{16.7.32}$$

Then Eq. (16.7.31) can be written

$$G = g + g^{-1} - 2 I \tag{16.7.33}$$

We also know that the basic orthonormals of $\underline{g}$, $\underline{g}^{-1}$, and $\underline{G}$ are the same and that, if the basic structure of $\underline{g}$ is

$$g = Q \delta Q' \tag{16.7.34}$$

then

$$g^{-1} = Q \delta^{-1} Q' \tag{16.7.35}$$

and

$$G = Q (\delta + \delta^{-1} - 2 I) Q' \tag{16.7.36}$$

It should also be noted that $\underline{g}^{-1}$ is the well known matrix of $\underline{n}$-2 order partial correlation coefficients.

16.7.4 The Optimal Error Covariance Matrix

Consider again the estimated covariance matrix

$$C_{W W} = R - 2 D_R^{-1} + D_R^{-1} R^{-1} D_R^{-1} \tag{16.7.37}$$

and the estimated error covariance matrix

$$C_{E E} = D_R^{-1} R^{-1} D_R^{-1} \tag{16.7.38}$$

We may wish to consider a scaling of $\underline{C_{W\,W}}$ in Eq. (16.7.37), and hence the same scaling of $\underline{C_{E\,E}}$ in Eq. (16.7.38), which will result in the best least square approximation to the identity matrix. We may then equally well consider the scaling of $\underline{R}^{-1}$ which is the best least square estimate of the identity matrix. Let

$$d\,R^{-1}\,d - I = \epsilon \tag{16.7.39}$$

and determine $\underline{d}$ so that

$$\mathrm{tr}\;\epsilon'\epsilon = \min = \psi \tag{16.7.40}$$

From Eqs. (16.7.39) and (16.7.40)

$$\psi = \mathrm{tr}\,(d\,R^{-1}\,d^2\,R^{-1}\,d - 2\,d\,R^{-1}\,d + I) \tag{16.7.41}$$

Let

$$R^{-1} = \rho \tag{16.7.42}$$

and from Eqs. (16.7.41) and (16.7.42)

$$\psi = \mathrm{tr}\,(d\,\rho\,d^2\,\rho\,d - 2\,d\,\rho\,d + I) \tag{16.7.43}$$

Let

$$V = d\,1 \tag{16.7.44}$$

From Eq. (16.7.44)

$$\mathrm{tr}\,(d\,\rho\,d) = V'\,D_\rho\,V \tag{16.7.45}$$

Let $\underline{\rho^{(2)}}$ be a matrix of the squared elements of $\underline{\rho}$. Then it can be shown that

$$\mathrm{tr}\ (d\ \rho\ d^2\ \rho\ d) = V'\ d\ \rho^{(2)}\ d\ V \tag{16.7.46}$$

From Eqs. (16.7.43), (16.7.44), (16.7.45), and (16.7.46)

$$\psi = V'\ d\ \rho^{(2)}\ d\ V - 2\ V'\ D_\rho\ V + n \tag{16.7.47}$$

Differentiating Eq. (16.7.47) symbolically with respect to $\underline{V'}$

$$\frac{\partial\ \psi}{\partial\ V'} = 4\ (d\ \rho^{(2)}\ d\ V - D_\rho\ V) \tag{16.7.48}$$

Equating Eq. (16.7.48) to zero

$$d\ \rho^{(2)}\ d\ V = D_\rho\ V \tag{16.7.49}$$

or from Eqs. (16.7.44) and (16.7.49)

$$\rho^{(2)}\ d^2\ 1 = D_\rho\ 1 \tag{16.7.50}$$

From Eq. (16.7.50)

$$d^2\ 1 = (\rho^{(2)})^{-1}\ D_\rho\ 1 \tag{16.7.51}$$

To scale $\underline{C_{W\ W}}$ in Eq. (16.7.37), therefore, so that the scaled $\underline{C_{E\ E}}$ part is the best least square approximation to an identity matrix, we write first from Eq. (16.7.37)

$$D_R^{-1}\ C_{W\ W}\ D_R^{-1} = D_R^{-1}\ R\ D_R^{-1} - 2\ D_R^{-1} + R^{-1} \tag{16.7.52}$$

Then letting $\underline{\gamma}$ be the scaled $\underline{C_{W\,W}}$ matrix, and using $\underline{d}$ given by Eq. (16.7.51), we have

$$\gamma = d\,(D_R^{-1}\,R\,D_R^{-1} - 2\,D_R^{-1} + R^{-1})\,d \qquad (16.7.53)$$